Investigating prognostic factors in Sepsis using Computational Intelligence methods

Universitat Politècnica de Catalunya

Carles Morales Boada

October 2016

Abstract

Sepsis is a major challenge in medicine. It is in fact a traversal condition affecting people of all ages and is not respectful of lifestyle choices. There are many diseases in medicine whose definition is uncertain and cause high rates of mortality in medical services, and *Sepsis* is a flagship example. In an intensive care unit (ICU), patients in an advanced stage of *Sepsis* carries a high burden, namely a high mortality rate (about 50%) and higher costs of treatment compared with other ICU patients.

After an operation, patients have a tendency to develop a phenomenon related to the mechanism of immune system. The pathophysiology of *Sepsis* in humans is poorly understood, even is one of the main causes of death for non-coronary ICU patients. It is a traversal condition affecting people of all ages. This syndrome was defined by consensus statement in 1992 to consist of certain criteria. Severe sepsis was defined as organ failure in the setting of sepsis, and septic shock was defined as severe sepsis where the organ failure was hypotension.

The ICU environment is one of the scenarios in which critical decision making tasks are most relevant with respect to the outcome for the patient, and it is in this specific context that this thesis tries to make a contribution.

The research reported in this document deals with the problem of *Sepsis* data analysis in general. On the one hand, a causal relationship study is made over a set of roughly one hundred different ICU measurements to detect hidden patterns among them, and to find the direct conditioners of the patients outcome, by means of probabilistic approaches (i.e. Bayesian Networks).

On the other hand, the problem of survival prediction in patients that have suffered *Septic Shock* is faced applying Computational Intelligence (i.e. Artificial Neural Networks) and Machine learning approaches (Support Vector Machines), and more specifically in two different approaches: (1) predicting directly the outcome of the patient and (2) predicting the organ dysfunction risk, that can also lead patient to death or severe cognitive consequences.

The data set used in this work is the public available MEDAN database, which consists of 412 patients that have suffered abdominal Septic Shock, of whom 201 died. The data were recorded from 71 German Intensive Care Units from 1998 to 2002.

Acknowledgements

I would start to thank to my thesis directors, with whom I am indebted, Alfredo Vellido and Vicent Ribas, for their disposal twenty-four hours a day to help and advise me in absolutely all questions and doubts that I had during the writing of this work, that were not just a few.

They introduced me to medical field, specifically to the study of Sepsis, which I was completely unaware, even of the name of the illness, but I did not doubt even a second to submerge myself in such interesting project, and even more motivated seeing the team that was disposed to guide me.

Again, I am very grateful to you and proud of the work done together.

I want also to thank to my family, specially mum and dad. They have done everything possible so that I can get here and present my master thesis. Somehow, this work is theirs.

Contents

Chapter 1

Introduction

Sepsis is a major challenge in medicine. It is in fact a traversal condition affecting people of all ages and is not respectful of lifestyle choices. Vulnerable groups such as patients with weakened immune systems, people with chronic diseases, new born babies, small children and the elderly are amongst the most at risk. Adhikari and colleagues estimated an incidence of sepsis of up to 19 million cases worldwide every year [1], being the true incidence presumably far higher. Many of those affected needlessly die or suffer permanent health issues due to lack of proper diagnosis and/or treatment. It is noteworthy to mention that this pathology is more common than heart attacks and claims more lives than any cancer. *Sepsis*, in fact, remains the primary cause of death from infection despite the manifold advances in modern medicine such as vaccines, antibiotics and critical care procedures. In the developing world, sepsis accounts for 60-80% of lost lives per year, affecting annually more than 6 million newborns and children. Furthermore, over 100.000 women contract sepsis in the course of pregnancy and childbirth [2].

In western countries, *Sepsis* or *Septic Shock* shows a prevalence of 3 cases per 1000 people, representing no less than the 25% of total Intensive Care Unit (ICU) admissions and accounting for around 1%-2% of all hospitalizations. Importantly, this pathology is associated to a very high mortality rate, which varies between 30% and 50% according to different studies [3, 4]. Massive resources are currently being invested in developing and evaluating potential therapies and considerable effort has been undertaken to understand the systemic inflammation and multiple-system organ failure that are characteristic of the severe version of sepsis [3].

Septic response can be portrayed as being one of the main contributing factors to around 750,000 cases per year only in the United States of America, whereof 200,000 end up dead, with a worrying increasing trend that is likely to rise annually by up to 8,5% [3], in contrast to what would be expected from a controlled management of such a widespread pathology.

There is also an important economic side linked to it, as it is associated to high costs of treatment in comparison with those of other ICU patients [5]; this fact has undoubtly reinforced the efforts towards achieving an understanding of sepsis, despite the current limitations in terms of prognostic estimation [6].

On top of this is the poor understandability of the patho-physiology of sepsis in humans. This syndrome was defined by consensus statement only in 1992 [7] to conform to certain criteria: The systemic inflammatory response syndrome can be self-limited or can progress to *Severe Sepsis* and *Septic Shock*. Along this continuum, circulatory abnormalities (intravascular volume depletion, peripheral vasodilatation, myocardial depression and increased metabolism) lead to an imbalance between systemic oxygen delivery and oxygen demand, resulting in global tissue hypoxia or shock [8]. The transition to serious illness occurs during the critical period commonly known as the "golden hours", when definitive recognition and treatment provide maximal benefit in terms of outcome.

Very recently, a new formal definition of *Sepsis* was agreed by consensus, downplaying the importance of SIRS as a key indicator of the pathology [9]. This redefinition is a key matter for this thesis.

It is not yet possible to accurately predict the behaviour of patients undergoing *Sepsis* due to the heterogeneity of patients groups and the variations in therapy strategies. Its prognostic

complexity is also aggravated by the challenging emergency situations in which most diagnoses are carried out, usually with several pathologies concurring at the same time. Hence, it is of paramount importance obtaining early indicators of the presence of *Sepsis* in order to allow doctors to act expediently at the onset of symptoms.

The ICU environment is one of the scenarios in which critical decision making tasks are most relevant with respect to the outcome for the patient, and it is in this specific context that this thesis tries to make a contribution.

Clinicians might benefit from at least partially automated computer-based decision support, and it is of crucial importance to design tools addressed to clinicians that are easily and rapidly interpretable. This thesis tries to contribute to solve the decision making challenges in this area by: analysing the indicators that are usually recorded in ICU patients and looking for strong causal relations among these indicators and the final outcome of the patient. This is made operational by defining a risk of death (*exitus*).

1.1 Artificial Intelligence in Medical Diagnosis

Artificial Intelligence (AI) has steadily found its way over the last decades in many real-world applications and the pace of this process has stepped up in the last decade. This is particularly true for applications in biology and medicine. As part of the application of AI methods in medicine and healthcare, medical diagnosis is one of of the most important aspects of the problem, and one with the most relevant human implications. Diagnosis, as any other human undertaking, is prone to error, and huge efforts have been made to assist experts in this decision making tasks.

Part of the uncertainty linked to medical diagnosis has to do with the vast amount of very diverse medical data available to the physician, which may become an excessive burden for human skills and experience. In this context, more-or-less automated data-based techniques for decision support, in the form of well-designed medical decision support systems (MDSS), may become essential in the task of extracting usable knowledge from complex medical data.

This is not enough to justify the use of AI as part of MDSS. We also need to address of model interpretability, otherwise risking the usability of our models, regardless their efficacy.

Analyzing patients' data and capturing all this knowledge in the form of computer models could help healthcare providers and even expert physicians to provide accurate advice to patients, and any practitioner could make use of that expertise whenever a patient case suggests the need for careful thought about some aspects of the pathology or the subsequent therapy. This is the core objective of Medical AI.

1.2 Motivation

As stated above, *Sepsis* is the result of the uncontrolled inflammatory response to an infection. At this stage it is also very important to note that, to the present day, *Sepsis* is a diagnose that can only be assessed with certainty *a posteriori* (i.e. when the condition has already taken place), but at the same time requires immediate action because of the rapid evolution to severe situation of this kind of disease and, whenever is possible, pre-emptively. It is also of great importance the rising trend that *Sepsis* is starring, specially in well-developed countries, faster than research in the field is evolving, because there are several factors that make early *Sepsis* prognosis a great medical and also technological challenge.

One of the most important is the "low exactitude" of the formal *Sepsis* definition. It has changed several times along the years, being the last update from 2016 [9], changing the weights of the key indicators in the overall definition that characterize the illness. Because of the weak robustness of the definition along the years makes even more important to apply the strength of Artificial Intelligence and computation in order to help defining more accurately the factors that activate *Sepsis*.

Another challenging characteristic is the environment where almost all the stages occurs: Intensive Care Units, where the diagnostic of the patient does not end in just a *Sepsis* clinical picture but usually combined with more pathologies, and an extremely data intensive environment, entailing an impossible challenge for the best clinician to take profit of all monitored indicators that are

continuously producing data from beat-to-beat (Blood Pressure, Heart Rate or ECG), hours (gas exchange, white blood cell count), to days (APACHE, SOFA, MODS).

On the basis of this scenario, it becomes extremely motivating to seize the opportunity to help in this important health issue, by means of the application of Artificial Intelligence knowledge. We are aware that this aim is not an easy road and, less still, fast.

1.3 Thesis Objectives

The main objective of this thesis is to obtain a better understanding about the pathophysiology of severe sepsis and in particular to assess the usefulness of the systemic inflammatory response syndrome (SIRS) in both outcomes (i.e. exitus letalis) and organ dysfunction. This is particularly important in the light of the new definitions of sepsis [9]. As secondary objectives we also expect to:

1. Improve our knowledge about the causal relationship between different clinical traits/variables and ICU outcome
2. Improve our knowledge about the causal relationship between different clinical traits/variables and organ dysfunction during sepsis.
3. Design a classification system for assessing both ICU outcome and organ dysfunction during the first hours of evolution.

1.4 Considerations about the Analysed Data set

This work uses the public available MEDAN database [10], which contains the data of 412 patients that suffered abdominal Septic Shock, of whom 201 died (48.8%). The data were recorded from 71 German Intensive Care Units from 1998 to 2002 by medical documentation staff. The data were transferred from paper to the database and, as such, typing errors are a common source of errors. The MEDAN database was originally composed by several tables, from which two were selected: Patient Information and Variables Measurements.

- *Patient information*: Contains non-temporal patient characteristics (i.e. height, age, sex, outcome,...)
- *Variable Measurements*: Contains temporal records of 103 different variables (i.e. blood pressure, leukocytes, pH, arterial, pO_2, SOFA,...), belonging to four kind of data: numeric, class, yes/no and score.

The measurements are related with a time stamp from when they were recorded. Taking into account that the table Variable Measurements is made up by more than 2.5 million recordings, the mean number of measures by each patient is about 6800 instances. It reflects the data-intensive environment that ICU is, and also both the necessity and the opportunity of taking profit of that amount of recorded data.

There is still not a lot of research done over this database, but it is mandatory to cite nice research done by their creators; Paetz *et al* [5] [11] [10] [12], and also from some more researchers, as Fialho *et al* [13] [14].

1.5 Expected Contributions

The main contribution of this thesis is to study the usefulness of SIRS in the continuum of sepsis and, therefore, in its definition through machine learning techniques. Moreover, it is also expected to detect a sub-set of variables that could be later used for assessing organ dysfunction at the onset of the syndrome with better sensitivity and specificity than current state-of-the-art methods.

1.6 Thesis Structure

This thesis is organized as follows:

- **Chapter 2** presents an overview of medical background in the *Sepsis* pathology. First of all, we provide information about the incidence of the illness in the world, and then follows a slight history of its definition and the changes it suffered during the last years.
- **Chapter 3** is devoted to the State of the Art of the *Sepsis* prognosis, preceded of a more detailed clinical definition of *Sepsis*, followed with the more important biochemical markers that are used in its prognosis and also the principal clinical scoring systems that are used to rank the patient gravity in relation with its clinical picture.
- **Chapter 4** presents the proper MEDAN database, with some demographics information. It follows with a gathering of the most common problems when dealing with clinical databases and the solution we took for them. Finally, the data preprocessing procedure is explained.
- **Chapter 5** is in charge of presenting the technical part Causal Discovery algorithms and its necessary background. Finally, the detailed implementation of Causal Explorer toolkit is explained.
- **Chapter 6** is mostly technical and contains a slight mathematical definitions of the two used methods for classification, that are Support Vector Machines (SVM) and Artificial Neural Networks (ANN).
- **Chapter 7** presents the Results and Discussion of the implementation of the different stages of this work: Firstly, the feature selection methodology to extract the desired subset of variables, then the results of applying Conditional Independence Maps to the patients information is presented, and finally the classification stage (predicting both patient outcome and organ dysfunction gravity) is explained.
- **Chapter 8** provides the conclusions of this master thesis and the main contributions.
- **Chapter 9** presents a publication derived from the thesis, focused in the application of Conditional Independence Maps.

Chapter 2

Medical Background: The Sepsis Pathology

As mentioned in the introduction, *Sepsis* is one of the main causes of death for non-coronary ICU patients. According to [6], it is the tenth most common cause of death. It remains the primary cause of death from infection, despite advances in modern medicine like vaccines, antibiotics, and intensive care. It is more common than heart attack, and claims more lives than any cancer, yet even in the most developed countries fewer than half of the adult population have heard of it. In the least developed countries, sepsis remains a leading cause of death.

Sepsis is a life-threatening medical condition that arises when the body's attempt to fight an infection results in the immune system damaging tissues and organs. This chaotic response, designed to protect us, causes widespread inflammation, leaky blood vessels, and abnormal blood clotting resulting in organ damage. Patients vary in their response; the severity of their sepsis and the speed at which it progresses is affected by their genetic characteristics and the presence of co-existing illnesses, as well as the numbers and virulence of the infecting micro-organism [15]. In severe cases, blood pressure drops, multiple organ failures ensue, and the patient may die rapidly from *Septic shock*.

Sepsis originally meant "putrefaction", a decomposition of organic matter by bacteria and fungi. Over time, an increasingly specific wide variety of definitions have been applied to sepsis, including sepsis syndrome, severe sepsis, septicemia, and septic shock [16, 17].

The definitions take into account the findings that sepsis may result from a multitude of infectious agents and microbial mediators and may not be associated with actual blood stream infection.

The term "Systemic Inflammatory Response Syndrome" (SIRS) was coined to describe the common systemic response to a wide variety of insults [18], later explained in detail. When SIRS is the result of a confirmed infectious process, it is termed sepsis. *Severe Sepsis* is defined as sepsis plus either organ dysfunction, or evidence of hypoperfusion or hypotension. *Septic shock* is a subset of *Severe Sepsis* and is defined as sepsis-induced hypotension, persisting despite adequate fluid resuscitation, with the presence of hypoperfusion abnormalities or organ dysfunction [19], as summarized in figure 2.1.

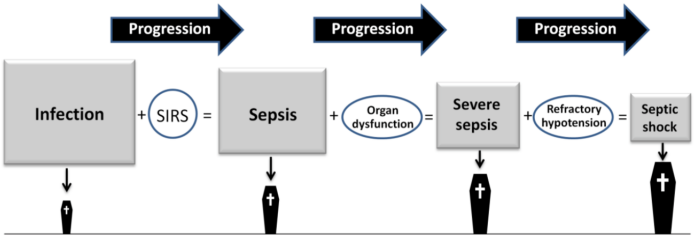


Figure 2.1: The progressive severity stages of sepsis. Increasing severity is characterized by lower incidence and increased case fatality rates. Adapted from Gille-Johnson et al [20].

Saving lives in this context depends not just on treatments that are specific to a particular infection, but rather on focusing on the early recognition and awareness of sepsis, rapid antimicrobial therapy and resuscitation, together with vital organ support. In short, sepsis is a medical emergency and, literally, each hour matters. A better understanding of sepsis as the final common pathway of illness due to infection is essential to drive any improvement on diagnosis, prognosis and treatment.

The specific statistics for the German population (the region where the dataset analyzed in this thesis belongs to) are similar to those of the developed world in general; the incidence of *Severe Sepsis* or *Septic Shock* can be estimated in 110 per 100,000 people, figures that are comparable with those of the incidence of acute myocardial infarction (143 per 100,000 people). With an estimated 40,000 deaths per year, *Severe Sepsis/ Septic Shock* is the third most frequent cause of death in Germany after coronary artery disease and acute myocardial infarction [21].

Figure 2.2 reflects a comparison of hospital admissions for stroke and myocardial infarction. It can be appreciated that, whereas these two remain stable over the same period, *Sepsis*, instead, presents a clear increasing trend.

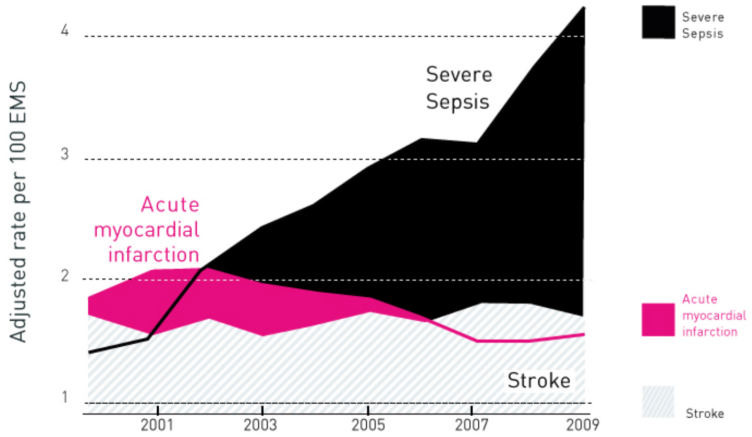


Figure 2.2: Temporal changes in the rates of hospitalizations with severe sepsis among Emergency Medical Services (EMS) encounters, adjusted for age, sex, and receiving hospital. A comparison with acute myocardial infarction (AMI) and stroke is provided for context. Adapted from Seymour et al. [22].

In Spain, the incidence of *Severe Sepsis* is quantified in around 104 cases per 100,000 people [23] and represent the most prevailing illness in the ICU. The impacting fact to reckon with is that its incidence grows by 7-9% annually, due to the longer life expectancy, the increase of use of body-invasive medical procedures, immunosuppression states caused by drugs, the treatment of patients with chemotherapy, etc.

These epidemiology statistics, even if restricted to the Spanish population, can again be extrapolated to other developed areas in the world [24].

The associated mortality rates vary depending on the stage of sepsis affecting the patient. So, *Sepsis* has 10%-20% of patients that die, while *Severe Sepsis* is 20%-50%, and *Septic Shock* is 40%-80% [24].

It is important to note that some patients that manage to survive may experience severe long-term cognitive decline following an episode of *Severe Sepsis*, but the absence of baseline neuropsychological data in most sepsis patients makes the incidence of this difficult to quantify or to study [25].

2.1 Evolution of Sepsis Definition

Sepsis has historically been seen as a condition that is difficult to identify and diagnose. As far back as 100 BC, Marcus Terentius Varro, the ancient Roman scholar and writer (116 BC–27 BC), was quoted as noting that *"small creatures, invisible to the eye, fill the atmosphere, and breathed*

through the nose cause dangerous diseases". Perhaps the most prescient description of sepsis was by the hand of the historian, philosopher, humanist and Renaissance author Niccolo Machiavelli (1469–1527), who stated that *"hectic fever, at its inception, is difficult to recognize but easy to treat; left unattended it becomes easy to recognize and difficult to treat"*. Although hectic fever is of course not the name by which we know *Sepsis* nowadays, the description of a disease that is difficult to recognize in its early stages, at a time when the condition may be easily treatable, and more difficult to treat in its later more obvious stages is a clear description of what we understand now as the more severe forms of *Sepsis*.

In an attempt to better clinically understand sepsis, a variety of definitions were posed during the past century. Among the earliest concepts was that which considered *Sepsis* as a systemic host response to an infection [26]. In fact, it was classically described by the American physician William Osler, that the patient appears to die from the body's response to an infection rather than from the infection itself. In 1972, this concept was reinforced in a medical review, noting that *"it is our response that makes the disease"* [27]. The general concept has been long considered a form of poisoning, often considered as blood poisoning, but more practically representing the presence of pathogenic organisms or their toxins in the blood or tissues.

It was the failure of these medical definitions, and after thousands of attempts at developing diagnostic tools to identify sepsis, that led to a consensus conference focusing on a way to clinically define sepsis. In 1992, the American College of Chest Physicians (ACCP) and the Society of Critical Care Medicine (SCCM) jointly published the consensus definitions of *Sepsis* [28]. These novel description of SIRS criteria and specific definitions for *Sepsis*, *Severe Sepsis*, *Septic Shock* and multiple organ dysfunction syndrome (MODS) were all critical developments in the investigation of the field (see fig 2.3). Since these consensus definitions had limitations in clinical use, they were revisited in 2001 in a new *Sepsis* Definition Consensus [29]. Although there were many limitations that were recognized, there was no better alternative identified. Significant consideration was given to the possibility of expanding the foundational "Systemic Inflammatory Response Syndrome" criteria to include other parameters that might be associated with *Sepsis*.

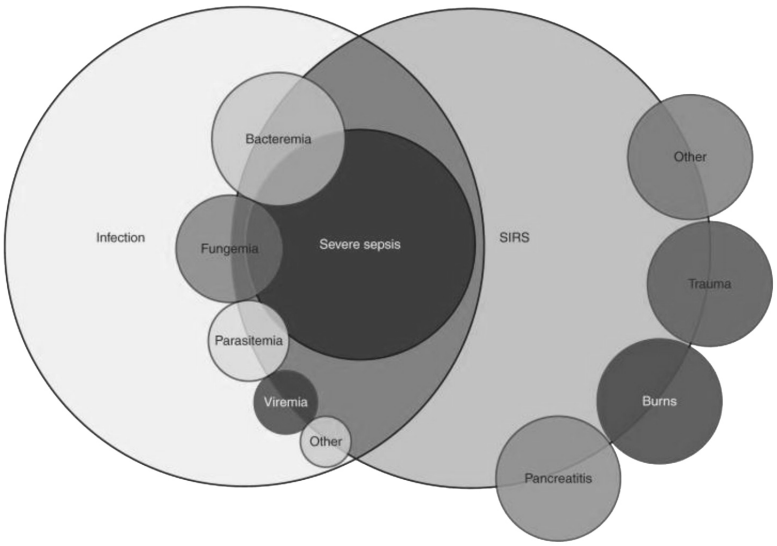


Figure 2.3: Relationship between systemic inflammatory response and infection, where the overlap indicates sepsis [24].

In addition, some of the criteria overlapped with the definitions developed for identifying organ dysfunction, which is a critical component of distinguishing *Severe Sepsis* and *Septic Shock*. Perhaps the most important result from the 2001 Consensus Conference was the proposal for a "Predisposition, Infection, Response and Organ dysfunction" (PIRO) system for staging sepsis. The concept of PIRO was analogous to staging cancer or other medical conditions, and it appears

that these criteria do allow for differentiating groups of patients with *Sepsis* [30].

The latest definition agreed by the major experts in *Sepsis* field took place in 2016 [9], in which the excessive focus on inflammation as a key indicator of *Sepsis* is questioned, as well as the inadequate specificity and sensitivity of the SIRS criteria. The redundancy of the term *Severe Sepsis* was finally agreed upon.

Chapter 3

Sepsis Prognosis: The State of the Art

The definition of *Sepsis* and its clinical manifestation have been, over the years, controversial issues, due both to the importance of the pathology, related to its high mortality rates, and to the difficulties found in the determination, in a consensual manner, of its definition in a way that allows standardizing patients' diagnosis and establishing a suitable path of treatment as soon as possible, preventing the evolution towards multi-organic failure, with far worse prognosis.

In 1992, the *American College of Chest Physicians* (ACCP) and the *Society of Critical Care Medicine* (SCCM) celebrated a Consensus Conference [31] where they agreed definitions in an attempt to overcome the limitations of the vague terminology that surrounded septic patients.

The SIRS concept was introduced as the generalized organism response against certain stimuli, the presence of whom can be produced by infectious or not infectious etiology, in combination with two or more of the following clinical symptoms:

- Temperature higher than 38°C or lower than 36°C.
- Heart rate higher than 90 beats per minute.
- Respiratory rate higher than 20 breaths a minute or $PaCO_2$ lower than 32 mmHg.
- Alteration in white blood cell count with more than 12.000 or less than 4.000/mm^3 or more than 10% of immature bands.

However, these criteria (more detailed in table 3.2) are not specific enough to diagnose the cause of the syndrome, or to identify different response patterns in patients. Thus, in a new Consensus Conference celebrated in 2001 [29], a wider list of diagnostic criteria of systemic inflammation in response to an infection was proposed. The identification of some of these criteria would lead us to state that a patient "seems septic". It is important to note that none of these criteria is specific for *Sepsis*.

3.1 Biochemical markers

Because these symptoms are not exclusive of *Sepsis*, the Consensus Conference of 2001 [29] confirmed the high sensitivity of clinic criteria that define SIRS, but also their low specificity, limiting their diagnostic usefulness and revealing the need to find out new bio-markers to improve its definition and a rapid and clear diagnostic.

A diversity of molecules have been studied as possible bio-markers of infection. The most relevant [32] include:

- **Procalcitonin (PCT)**: PCT levels lower than 0,5 ng/L indicate viral infections or non-infectious chronic inflammatory processes. Values between 0,5 and 2 ng/L are given in patients with multiple traumas, burned or postsurgical [33]. Concentrations higher than 10

ng/L occur in patients with sepsis, being levels higher than 10 ng/L from patients of *Severe Sepsis* or *Septic Shock*. PCT production is induced by bacterian endotoxines and exotoxines and several cytokines. It appears in plasma 3 hours later from the beginning of *Sepsis* and reachs a serum "peak" at 6 hours, remaining up to 24 hours. Specificity and sensibility of PCT as *Sepsis* indicator is clearly higher than other biomarkers [29]. A 2013 review concluded moderate-quality evidence exists to support use of the PCT level as a method to distinguish sepsis from non-infectious causes of SIRS [34], the same review found the test's sensitivity to be 77% and the specificity to be 79%. The authors suggested PCT may serve as a helpful diagnostic marker for sepsis, but cautioned that its level alone cannot definitively make the diagnosis.

- **C-reactive protein (CRP)**: Hepatic synthesis protein is an unspecific inflammation marker. Standard values are lower than 10 mg/L. Povoa, in an study realized in critical patients [35] found that a threshold value of 50 mg/L discriminates the infectious origin of inflammatory response with respect to other ones, with a sensitivity and specificity of 98% and 75% respectively. However, no significance differences were found in CPR levels in patients with *Sepsis*, *Severe Sepsis* or *Septic Shock*.

- **Neutrophil elastase (EN)**: This protein is released in the plasma in SIRS scenarios, but it is not a good indicator to differentiate SIRS from sepsis. Is secreted by neutrophils and macrophages during inflammation, it destroys bacteria and host tissue [36]. Important increases of elastase in blood are correlated with high mortality in septic patients [37].

Biochemical markers of *Sepsis* can only be used in clinical practice if they meet certain criteria directly related to therapeutic changes. Currently, only PCT meets these requirements: bacterial infection diagnosis, SIRS severity diagnosis and progression of infection to *Sepsis* and from *Severe Sepsis* to *Septic Shock*, as well as the response to the therapy.

It is important to note that in our studied database only one out of the three most explicative bio-markers to detect an earlier infection that can lead to SIRS diagnosis is available, namely the C-Reactive Protein. The current problem is the way to profit from all this knowledge, because the clinical mechanisms, and specially in ICU environments, make difficult to acquire biologic measures in patients in a rapid and frequent-enough manner, and it is more profitable to base the studies on more physiological or symptomatic oriented measures.

The increasing prevalence of this pathology even in developed countries and the high mortality rates that entails justify the need for a quantitative approach to predict mortality due to sepsis in the ICU. The extreme demands of this clinical environment further require prediction methods that are both robust and feasible within the constraints of an ICU.

The rationale for using scoring systems in a clinical environment is to ensure that the increasing complexity of any disease in treated patients is consistently represented for all those involved in the form of evaluations and descriptions. A specific goal of severity scoring systems is to use these important patient attributes to describe the relative risks and to identify where along the continuum of severity the patient resides. This should reduce the variability due to patient factors so that the impact or reactions of patients to new or existing therapies can be more precisely determined. Also, more precise measurements of patient risk should lead to new insights into disease processes and serve as a tool with which clinicians could more accurately monitor patients and implement the use of new therapies [38].

3.2 Clinical Scoring Systems

In this section we define the main concepts that surround *Sepsis* pathology.

Systemic Inflammatory Response Syndrome

As stated above, *Sepsis* is defined as the "*systemic response to infection*". It is apparent that a similar, or even identical response can arise in the absence of any infection. Therefore, the term SIRS is proposed to describe this inflammatory process, independent of its cause.

The innate ability of the body to defend itself is based on three elements: external barriers against invasion and tissue injury, non-specific systems against foreign pathogens and debris, and anti-specific responses to foreign pathogens [39]. Inflammation is the body's initial non-specific response to tissue injury produced by mechanical, chemical or microbial stimuli. Inflammation is a rapid highly amplified controlled humoral and cellular response: the complement, kinin, coagulation and fibrionolytic cascades, are triggered in tandem with the activation of phagocytes and endothelial cells [40].

SIRS was developed to imply a clinical response arising from a non-specific insult and includes two or more defined variables below.

1. Body temperature higher than 38°C or lower than 36°C.
2. Heart rate higher than 90 beats per minute (bpm).
3. Tachypnea, manifested by a respiratory rate higher than 20 breaths per minute or hyperventilation indicated by $PaCo_2$ of less than 32 mmHg.
4. Alteration in the white blood cell count, such as a count higher than 12,000/cu mm or lower than 4,000/cu mm, or the presence of more than 10% immature neutrophils.

Sepsis is defined as SIRS with a documented infection, and the sequelae of SIRS/sepsis is MODS, which can be defined as failure to maintain homeostasis without intervention.

Sequential Organ Failure Assessment Score (SOFA)

In 1994, the European Society of Intensive Care Medicine (ESICM) organised a consensus meeting to create the SOFA score [41], with the aim of objectively and quantitatively describing the degree of organ dysfunction or failure over time.

It was created to improve the understanding of the natural history of organ dysfunction/failure and the interrelation between the failure of various organs and systems, and also to assess the effect of new therapies on the course of organ dysfunction/failure to characterise patients at admission into the ICU or evaluate treatment efficiency.

Therefore, SOFA was not originally created to be used as a mortality predictor, but to describe complications of the critically ill. However, several articles can be found (see [42]) that prove that a SOFA score greater than 7 has important ICU outcome prediction capabilities. Moreover, the SOFA score in combination with additional parameters provides a very powerful set of predictors not only to predict patients outcome but also to detect changes in *Sepsis* severity.

SOFA limits the number of organs/systems investigated to six: respiratory (inspiration air pressure), coagulation (platelet count), liver (bilirubin), cardiovascular (hypotension), central nervous system (Glasgow coma score), and renal (Creatinine or urine output). The scoring for each organ/system ranges from 0 to 4, being 0 for normal function and 4 for maximum failure/dysfunction. The final SOFA score is the addition of all six indexes and, therefore, the maximum possible SOFA score is 24, corresponding to maximum failure for all six organs/systems. A more detailed definition of SOFA score can be found in table 3.1.

In the light of what has been described so far and from a practical point of view, a SOFA score greater than 1 corresponds to MODS, while Cardiovascular SOFA scores greater than 2 correspond to *Septic Shock*. Normally, SOFA scores are calculated at ICU admission. However, daily calculations of SOFA scores (Dynamic SOFA [43]) provide valuable information about organ dysfunction evolution and prognosis. Our database contains one daily measurement of SOFA score for almost all patients, which represents a rich source of quality information.

Simplified Acute Physiology Score (SAPS)

Since 1981, several severity scores have been proposed for ICU patients. The first ones were acute physiology and chronic health evaluation (APACHE, APACHE II), Simplified Acute Physiology Score (SAPS); later, Mortality Probability Model (MPM) [45] and APACHE III were introduced. In 1993, Le Gall *et al.* [46] described the SAPS II scoring system, to develop a method for converting the score to a probability of hospital mortality.

Table 3.1: SOFA Score table adapted from [41]. MAP stands for Mean Arterial Pressure, DPM for dopamine, DBT for dobutamine, AD for adrenaline, and NAD for Noradrenaline. Dosages are given in $\mu/Kg\Delta min$

SOFA Score Points	**1**	**2**	**3**	**4**
Respiration PaO_2 / FiO_2 mmHg	<400	<300	<200	<100
Coagulation Platelet Count: Platelets x $10^3/mm^3$	<150	<100	<50	<20
Liver Bilirubine [mg/dL]	1.2 - 1.9	2.0 - 5.9	6.0 - 11.9	>12
Cardiovascular Hypotension	MAP <70	DPM or DBT <= 5	DPM >5 AD <0.1 NAD <0.1	DPM >15 AD >0.1 NAD >0.1
Central Nervous System	13 - 14	10 - 12	6 - 9	<6
Renal Creatinine [mg/dL] or Urine Output	1.2 - 1.9	2.0 - 3.4	3.5 - 4.9 or <500 ml/day	>5 >200 ml/day

The original SAPS index was defined in 1984 and uses 14 routinely measured biologic and clinical variables [47] to develop a scoring system to calculate the risk-of-death (ROD) in ICU patients. Each variable has assigned a value that ranges from 0 to 4. (So the score ranges from 0 to 4 x 14 = 56).

As aforementioned, the definition was updated in 1995 and called SAPS II [46]. The SAPS II score is made up of 17 variables – 12 physiological variables, age, type of admission (scheduled surgical, unscheduled surgical, or medical), and 3 variables related to the underlying disease: acquired immunodeficiency syndrome, metastatic cancer, or haematological malignancy. Table I presents the definitions of variables constituting the SAPS II scoring system. Points assigned for each variable vary from 0-to-3 (for temperature) to 0-to-26 for the Glasgow coma scale.

It has been reported that SAPS presents a sensitivity and specificity of 0.69 for a cut-off value of 12 [47] and a population with more pathologies in patients than sepsis. Further studies [48] have been made analysing the usefulness of SAPS II in mortality prediction in ICU patients and the results are relevant. Similarly, studies about SAPS score have been carried out in [49, 50].

Several studies led by Drs. Ribas and Vellido at UPC treated *Sepsis* prediction by means of, among others, SAPS and SOFA scores [51, 52, 42, 51].

Acute Physiology and Chronic Health Evaluation II (APACHE II)

APACHE II is a severity-of-disease classification system. It uses a point score based upon initial values of 12 routine physiologic measurements, age, and previous health status to provide a general measure of severity of disease [53]. After admission to an ICU, an integer score from 0 to 71 is computed for the patient on the basis of several measurements. Higher scores imply a more severe disease and, therefore, a higher ROD.

APACHE II was designed to measure the severity of disease for adult patients admitted to ICUs. The minimum age is not specified in the original study [53], but it is commonly recommended to use APACHE II only for patients older than 15 years. This scoring system is applied in in different ways.

- Some procedures are only carried out in, and some drugs are only prescribed to, patients with a given APACHE II score.
- The APACHE II score can be used to describe the morbidity of a patient when comparing their outcomes with that of other patients.
- Predicted mortalities are averaged for groups of patients in order to specify the group's morbidity.

3.3 Current state in Sepsis Prognosis

The treatment of sepsis has placed a serious burden on health care systems, with an estimated US $14.6 billion spent annually on hospitalizations only in the USA [54]. Progression from *Sepsis* to *Severe Sepsis* is associated with increased mortality and morbidity, including permanent organ damage, cognitive impairment, and physical disability. Provision of appropriate treatment early in the development of sepsis has been associated with improved patient outcomes.

The benefit of these interventions is dependent upon the underlying short-term risk for mortality of the patient. For patients at high short-term risk, aggressive treatment and broad-spectrum antibiotics significantly decrease their mortality risk [55]. However, for low-risk patients, the associated risk of aggressive treatment outweighs their benefit. For this reason, it is of paramount importance to rapidly and accurately stratify patients with sepsis according to the predicted risk at the onset of the syndrome [56].

Previous studies have demonstrated that machine learning methods can be incorporated into electronic health records (EHRs) to predict clinically relevant outcomes in patients with sepsis.

In 2010, Peelen et al [57] developed a dynamic Bayesian Network (BN), modelling the progression of organ failure based on the Sequential Organ Failure Assessment (SOFA) score in the intensive care unit (ICU) [41].

Additionally, artificial neural networks (ANN) have been demonstrated to have important capabilities in diagnosis and prediction in medical field [58]. Specific studies predicting critical states of sepsis, where a critical state is defined as a physiologic state proximally associated with a worsening clinical condition or death, can be consulted in [59].

Progression from sepsis to severe sepsis was found to be accurately classified by Support Vector Machines (SVMs) [60, 61, 52, 62]. Moreover, dynamic BNs can predict mortality in sepsis and ICU patients, as demonstrated in [63, 64].

Also some studies of sepsis mortality prediction using decision trees [51] and graphical models [65] in ICU patients treated with statins have been recently published reporting promising results.

Although the aforementioned studies demonstrate the feasibility of extracting clinically relevant information pertaining to patients with sepsis, they do not specifically deal with the early identification of sepsis. Early detection of sepsis is challenging as the infection is not always clinically evident.

In addition, the signs that constitute the SIRS criteria were selected to be sensitive, but not necessarily specific for sepsis, making early diagnosis of the syndrome prone to false-positive classifications, as our experiments show.

On the contrary, the scoring systems previously explained (SOFA, SAPS, APACHE) have shown reasonable discrimination when classifying patients with sepsis admitted to the ICU into groups with high-risk and low-risk mortality. Although some studies have addressed the problem of sepsis patients' stratification according to risk before ICU care is needed [66], further investigation in classification systems that can accurately assess risk is still necessary .

Sepsis (documented or suspected infection plus ≥1 of the following)†
General variables
Fever (core temperature, >38.3°C)
Hypothermia (core temperature, <36°C)
Elevated heart rate (>90 beats per min or >2 SD above the upper limit of the normal range for age)
Tachypnea
Altered mental status
Substantial edema or positive fluid balance (>20 ml/kg of body weight over a 24-hr period)
Hyperglycemia (plasma glucose, >120 mg/dl [6.7 mmol/liter]) in the absence of diabetes
Inflammatory variables
Leukocytosis (white-cell count, >12,000/mm^3)
Leukopenia (white-cell count, <4000/mm^3)
Normal white-cell count with >10% immature forms
Elevated plasma C-reactive protein (>2 SD above the upper limit of the normal range)
Elevated plasma procalcitonin (>2 SD above the upper limit of the normal range)
Hemodynamic variables
Arterial hypotension (systolic pressure, <90 mm Hg; mean arterial pressure, <70 mm Hg; or decrease in systolic pressure of >40 mm Hg in adults or to >2 SD below the lower limit of the normal range for age)
Elevated mixed venous oxygen saturation (>70%)‡
Elevated cardiac index (>3.5 liters/min/square meter of body-surface area)§
Organ-dysfunction variables
Arterial hypoxemia (ratio of the partial pressure of arterial oxygen to the fraction of inspired oxygen, <300)
Acute oliguria (urine output, <0.5 ml/kg/hr or 45 ml/hr for at least 2 hr)
Increase in creatinine level of >0.5 mg/dl (>44 μmol/liter)
Coagulation abnormalities (international normalized ratio, >1.5; or activated partial-thromboplastin time, >60 sec)
Paralytic ileus (absence of bowel sounds)
Thrombocytopenia (platelet count, <100,000/mm^3)
Hyperbilirubinemia (plasma total bilirubin, >4 mg/dl [68 μmol/liter])
Tissue-perfusion variables
Hyperlactatemia (lactate, >1 mmol/liter)
Decreased capillary refill or mottling
Severe sepsis (sepsis plus organ dysfunction)
Septic shock (sepsis plus either hypotension [refractory to intravenous fluids] or hyperlactatemia)¶

† In children, diagnostic criteria for sepsis are signs and symptoms of inflammation plus infection with hyperthermia or hypothermia (rectal temperature, >38.5°C or <35°C, respectively), tachycardia (may be absent with hypothermia), and at least one of the following indications of altered organ function: altered mental status, hypoxemia, increased serum lactate level, or bounding pulses.
‡ A mixed venous oxygen saturation level of more than 70% is normal in newborns and children (pediatric range, 75 to 80%).
§ A cardiac index ranging from 3.5 to 5.5 liters per minute per square meter is normal in children.
¶ Refractory hypotension is defined as either persistent hypotension or a requirement for vasopressors after the administration of an intravenous fluid bolus.

Table 3.2: Diagnostic Criteria for Sepsis, Severe Sepsis and Septic Shock, adapted from Angus et al [44]

Chapter 4

Data Preprocessing: Preparing MEDAN database

Although it might seem surprising to devote a separate chapter to discuss the preprocessing of the MEDAN database used in this study, the fact is that its unique characteristics make it necessary to stress that a sizable proportion of the research reported in this thesis had in fact to do with the preparation of the data so that they could be further analyzed for knowledge extraction. It is fair to say, in any case, that this is not unusual in real world applications.

4.1 Dataset information

This work uses the publicly available MEDAN database [10], which contains data corresponding to 412 patients, 243 of them males and 169 females, with abdominal septic shock. Out of these, 201 died (a 48.8%), as shown in figure 4.1.

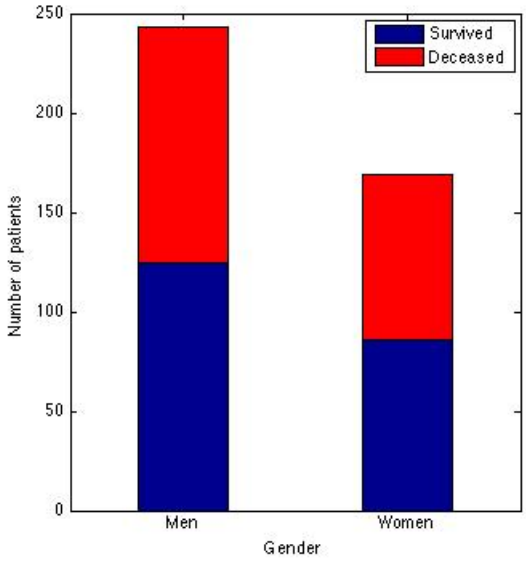


Figure 4.1: Ratios of deceased and surviving patients by gender

The data were recorded from the ICUs of 71 German hospitals, from 1998 to 2002, by medical documentation staff. This process was carried out by hand. The data were digitalized from paper to an electronic database and, in the process, typing errors became a common source of database quality issues [11]. It must be stressed, though, that the broad scale of the study (collecting data

from so many different hospitals is an unusual procedure) make this limitations still woth the effort from a data analysis viewpoint.

The Medan database was originally composed by several tables, including information about patients profiles, diagnoses, delivered medication, surgeries applied, eyc. In our work, we selected:

- **Patient information**: which includes physical characteristics of patients (age, weight, height, gender,...). It includes 103 different variables, which are numeric (real-valued or scores), categorical and binary.
- **Patient measures**: which consist of a large table with all measures and recordings of patients in an unsorted manner, identified by a unique code for each patient and each measured variable. Each measure is paired with a date and time vector.

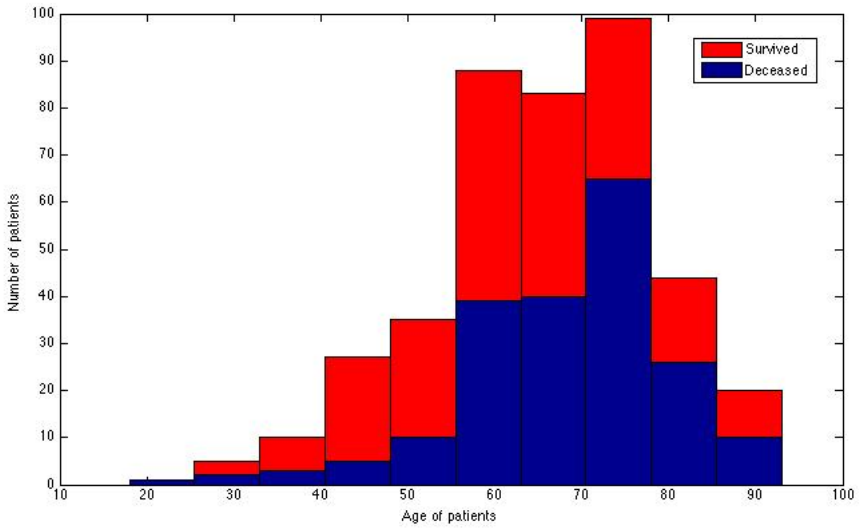


Figure 4.2: Histogram of attended patients by age, separated according to their final output.

4.2 Problems in Clinical Databases

As we know, databases are a set of records arranged in a regular structure that facilitates their management and any subsequent analytic treatment, but they tend to be incomplete and contain inaccurate data. This is the main reason why it is important applying some **data preprocessing processes** to ensure eliminating those factors that might lead us to reach wrong and distorted conclusions [58].

- *Different measured variables for each patient*: The gathering of information coming from different hospitals entails different variables having been measured in different ICUs, which adds a new difficulty to handle at the data gathering stage and, more importantly, a new difficulty to profit from variables that are only present in a small subset of the whole data. This situation yields many variables with a high percentage of missing values.
- *Different length of stay of patients at the ICU*: In clinical environments in general and more specifically in our case of study, patients remain in the ICU for different lengths of stay. This implies that, for some patients, there are recorded data for one or a couple of days, whereas, for others, there are measures spanning a whole week. Thus, in the overlapping and synchronising stage of the recordings, there is an important lack of logs for the main part of patients in the last days of the longest periods of stay.

  For this reason we decided to focus on three crucial moments for Sepsis identification and treatment: (1) time of admission to the ICU, (2) 24 hours after admission and (3) 48 hours

later. When no specific information about the times of admission/discharge of the patients to/from ICU were available, the starting and ending times were defined as the earlier and later measures registered for that patient in the database.

- *Measurements are carried out at different times of the day with a different frequency*: Among the whole set of different variables that are commonly measured in all hospitals, there are some with a recording frequency of minutes (e.g., heart rate or blood pressure), some with more of a few hours (e.g., temperature, central venous pressure (CVP), insuline), and some with a frequency close to a day (i.e. SOFA, APACHE II). Apart from the differences in the average sampling periods, there is a lot of variation of measured time stamps within records of the same variable. It adds the further difficulty of handling the registers in order to make them comparable and more or less homogeneous.

- *Many variables have a high percentage of missing data*: This problem goes hand in hand with the diversity of measured variables among different hospital ICUs. A variable that is only recorded in an specific hospital involves a huge number of missing values for that variable in the rest of the patients that were not treated in that hospital. Feature selection processes will thus be of paramount importance to face this problem.

- *Feature variables should be selected*, in order to avoid the so called "Curse of Dimensionality" effect. Aside from the "information holes" that add the consideration of seldom used variables, as explained in the previous point, a rigorous selection of the best subset of variables to take into account in further steps on the analyses is crucial.

Nevertheless, the preprocessing approach proposed in this work led to achieve better modelling results.

4.3 Preprocessing Procedure

The data preprocessing work was focused on obtaining three different subsets, each of them containing the bigger possible number of measures of the variables belonging to an specific moment of the ICU stay: (1) admission moment, (2) 24 hours and finally (3) 48 hours after admission. Hence, we deal separately with three subsets that contain just one measure (at most) of each of the 103 variables.

In order to select the proper sample measures that fitted better into some of the three subsets, when the algorithm does not find a value in that expected time position, it looks for the closest-in-time recorded value [67]. If there is no value around a requested time position, it is filled by *NaN* "flag values". In figure 4.3, the distribution of patients with respect to their related available data can be appreciated.

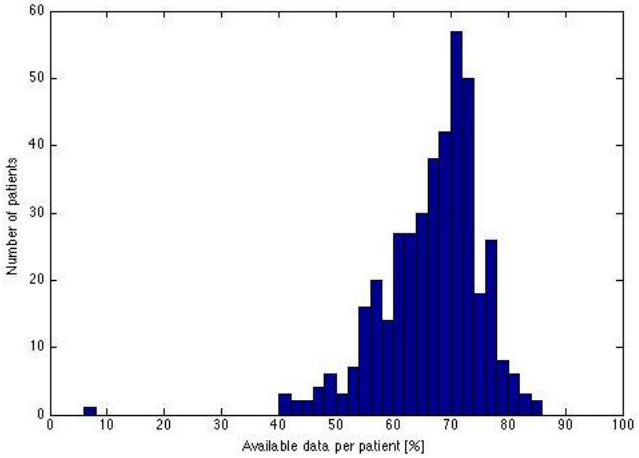


Figure 4.3: Available data per patient

Once the large original data set got reduced to three smaller subsets of size 412 x 103, we selected both those patients and variables with a small enough percentage of missing data, which guarantees the quality of the information used subsequently for modelling. The selection of patients and variables was based on the following steps, oriented so as to deal with the common problems in clinical databases explained in the previous section:

1. **Discard variables with an average sampling period bigger than 24 hours**. (Detailed information of variables and average sampling time roughly lower than a day can be found in Table 4.1). The result is a reduction of the data set to 74 variables.

2. **Discard variables with a missing data rate higher than 20%**. The decision of cutting at the 20% threshold is strict but fair, in order to retain the consistency of the data. From the previous selection of 74 variables, this step further reduced the data set to 52 variables.

3. **Discard patients with a missing data rate higher than 20%**. The previous steps already greatly reduced the existing number of missing values so that only remaining patient with an availability of data below 80% had to be removed.

The final data set, ready to be fed to the Causal Discovery Algorithms consisted of 411 instances and 52 variables: still of significant size for analysis.

Once the desired shape and size was given to the data set, the next step consisted in studying the probabilistic and causal relations between the variables, with the purpose of discovering potentially strong links. We made use of Causal Discovery algorithms and, more particularly, of Causal Independence Maps, as explained in the following chapter.

Variable	**frequency (h)**
Heart rate	1,2
Systolic blood pressure	1,2
Diastolic blood pressure	1,22
Mean blood pressure	1,37
Norepinephrine Perfusor	1,47
Ventilation mode	1,54
FiO_2	1,59
Dopamine Perfusor	1,63
Adrenaline syringe pump	1,69
Dobutamine Perfusor	1,73
Dobutamin single dose	1,81
O_2Gabe	2,03
PEEP	2,31
Dopamine single dose	2,34
Urine	2,59
IE	2,66
Temperature	2,95
HCLLoesung	3,12
Sedative narcotics	3,53
O_2 Saturation	3,66
Insulin	3,9
Analgesics	4,08
Ventilation rate	4,08
Pentaglobin	4,17
CVP	6,03
Hemofiltration	6,93
THAMTris	7
Antibiotics	7,05
Broncholyse	7,68
PH	7,73
PCO_2 arterial	7,78

Anticoagulation	7,88
Peak	7,95
PO_2 arterial	8,02
Base Excess	8,08
Bicarbonate	8,39
Antihypertensives	8,59
Blood sugar	8,93
GCSF	9,03
Immunosuppression (cor...	9,31
Loop diuretics	9,46
Parenteral nutrition	9,95
PCWP	10,44
Enteral nutrition	11,08
Hemoglobin	11,97
Crystalloids	12
Potassium serum	12,36
Antihypotensives	13,2
Sodium serum	13,41
Antiarrhythmic	13,41
Hematocrit	13,85
PPSB	15,83
Leukocytes	16,1
Platelets	16,31
Antifungals	16,39
Erythrocytes	16,56
Albumin5	16,59
PTT	18,42
TPZ	18,95
Sodium bicarbonate	19,03
Immunoglobulins Polyglob	19,41
FFP	20,47
Creatinine serum	20,81
Catheter	21,12
Diuretics	21,24
Norepinephrine single dose	22,15
SIRS	23,12
Urea	23,63
Dialysis	23,64
SOFA	24,34
Apache2	24,34
SAPS2	24,34
MODS	24,34

Table 4.1: Sorted average of sampling periods of the variables without any preprocessing or resampling, until 24h, computed by data of all patients for which each variable was measured

Chapter 5

Causal Discovery

Discovery of causal knowledge is crucial for advancing research in general and, very specifically, in medical research. Medical experts need to know the factors that cause a disease to devise new therapeutic procedures.

Classical statisticians often quote "*association is not causation*" to indicate that causal discovery is impossible without more evidence than that provided by frequentist experiments. As an illustrative example, simply observing a high occurrence of yellow stains on the fingers of patients that are suffering lung cancer with respect to normal subjects does not imply a causal relation between cancer and staining, because in reality heavy smoking is causing both to co-occur often.

In this chapter, we describe the causal discovery algorithmic methods applied in the research reported in the thesis, namely Causal Probabilistic Networks (CPN). The relevance of Causal graphical models such as CPNs are recognized in computational biology, biomedicine and bioinformatics as relevant representations capable of modeling causal relationships more precisely than standard clustering or regression models. They also have sound statistical foundations for inferential modeling [68] and for handling noise and missing data.

Unfortunately, discovering causal relations by randomized experimentation maybe impractical, unethical, or simply impossible. Recent advances in computational causal discovery theory and algorithm research and development mathematically prove the feasibility of causal discovery from observational data [69, 70, 71].

A classical study [72] carried out back in 1971 showed these basic facts in the medical area. At the university clinic of Leeds, United Kingdom, 472 patients with acute abdominal pain were examined and diagnosed. With simple, probability-based methods (Bayes classification) the diagnostic decision probabilities were computed from a data base of 600 patients. Additionally, a second set of probabilities were computed by using a synthetic data base of patients built from the interviews of experts and questionnaire sheets about *typical* symptoms. Then, the 472 cases were diagnosed by an expert round of three experienced and three junior physicians. The results of this experiment was as follows:

- Best human diagnosis (most experienced physician): 79.7%
- Computer with expert data base: 82.2 %
- Computer with 600 patients data: 91.1 %

The conclusion is clear: humans cannot *ad hoc* analyze complex data without errors [58], and here is where Causal Probabilistic analysis in medical environments makes real sense.

5.1 Causal Probabilistic Networks (CPN)

Causal probabilistic Networks (CPN), (a.k.a. Bayesian Networks (BN) or Belief Networks) are computational and mathematical objects that compactly represent joint probabilities by means of a directed acyclic graph (DAG) denoting dependencies and independencies among variables and conditional probability distributions for each variable given its parents in the graph [73].

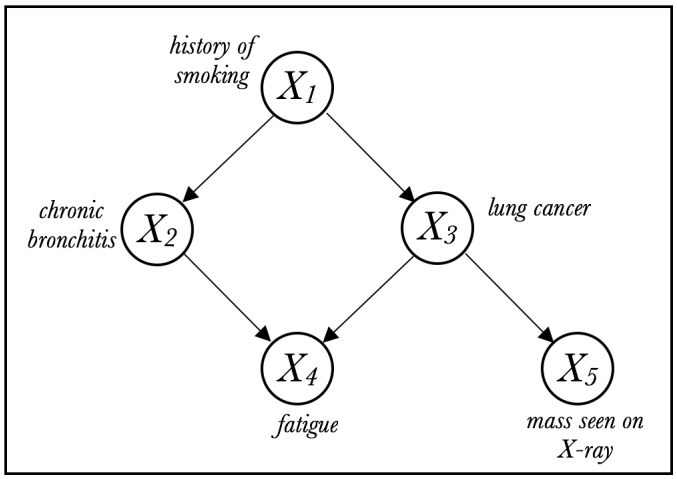


Figure 5.1: A hypothetical Bayesian Network structure

Probabilistic Networks are graphical models of (causal) interactions among a set of variables, where the variables are represented as nodes of a graph and the interactions (direct dependencies) as directed links between the nodes. Any pair of unconnected/nonadjacent nodes of such a graph indicates (conditional) independence between the variables represented by these nodes under particular circumstances that can easily be read from the graph. Hence, probabilistic networks capture a set of (conditional) dependence and independence properties associated with the variables represented in the network.

In the simplest case, a BN is specified by an expert and is then used to perform inference. In other applications, the task of defining the network is too complex for humans. In our case, we used BNs to determine network structure and the parameters of the local distributions, which must be learned from data, and for that we used the publicly available *Causal Explorer* toolbox [74], described in the next section.

The fundamental axiom of BNs is the *Markov Condition* (MC) that allows for a concise factorization of the joint distribution and captures the main characteristic of causation in macroscopic systems, namely that causation is *local*.

5.1.1 The Markov Condition

The independence relationships represented by the structure of a BN model are given by the Markov Condition: any node in a Bayesian network is conditionally independent of its non-descendants, given its parents.

A descendant of a node X is a node Y that can be reached by a directed path from X to Y. The Markov condition, as stated above, permits the factorization of a joint probability distribution on model variables $X_1, X_2, ..., X_n$ into the following product [75]:

$$P(X_1, X_2, ..., X : n) = \prod_{i=1}^{n} P(X_i \mid parents(X_i)) \tag{5.1}$$

where $parents(X_i)$ denotes the set of nodes with arcs into X_i. If X_i has no parents, then the set $parents(X_i)$ is empty, and therefore $P(X_i \mid parents(X_i))$ is just $P(X_i)$.

In the BN example shown in Figure 5.1, the application of equation 5.1 allows the derivation of a joint probability on the five model variable as follows:

$$P(X_1, X_2, X_3, X_4, X_5) = P(X_5 \mid X_3)P(X_4 \mid X_2, X_3)P(X_3 \mid X_1)P(X_2 \mid X_1)P(X_1) \tag{5.2}$$

Let us consider a node X in a BN. The *Markov Blanket* X is defined as the set of nodes consisting of the parents of X, the children of X, and the parents of the children of X. From the Markov Condition follows that if we condition on the values of each node in the Markov Blanket of X, then X is probabilistically independent of all other nodes in the network other than X and

its Markov Blanket. More detailed definitions of BNs and CPNs are out of the scope of this work; for further information see, for example, [75, 76, 77].

We used the Markov Blanket statement in order to select the nodes (i.e. variables) that impact more in the predicted variable: the output of the patients.

5.2 Causal Explorer

The implementation in our study is based on the *Causal Explorer* public library [74], which allows studying both causal relationships and variable selection.

The appeal of CPNs it that, contrary to the heuristic approaches for generation of causal hypotheses in bioinformatics and biomedical research, the recently developed theory of causal induction using graphical models and related distributions, provides guarantees for highly sensitive and specific discovery of causal relationships [70].

For the experiments reported in the next chapter, concerning the study of the causal relations among the patients' data records and their final output, we used the three different data subsets described in previous chapters as such, as well as narrowed by means of using only the ten and twenty more commonly available variables in patients' clinical records.

The rationale for this selection is, basically, trying to approximate the experiment to clinical reality constraints and to facilitate its visual comprehensibility. The more often-recorded variables are likely to be the ones that are considered by the medical experts to be relevant in regular examinations at the ICU, for that reason, we focus on analyzing the predictive power of these most available inputs.

Finally, we also processed the graphs in the same situations but excluding the final outcome variable (deceased or not) from the data as to unravel the hidden behaviours among the rest of measures that might have been eclipsed by the outcome variable.

A further step was to repeat the same procedure but reducing the "data availability" constraint of 20 and 10 variables to a maximum percentage of missing data, which we defined to be 20%; still a conservative decision to assure data trust.

In our study we use Causal Explorer with the three-phase dependency analysis algorithm (TPDA). The TPDA algorithm consists of three phases: drafting, thickening and thinning. In the drafting phase, TPDA produces an initial set of edges based on a simpler test (basically just having sufficient pairwise mutual information). This first draft is a graph without loops.

In the second phase, TPDA adds edges to the current graph when the pairs of nodes cannot be separated using a set of conditional independence tests. The graph produced by this phase will contain all the edges of the underlying dependency model. In the thinning phase each edge is examined and it will be removed if the two nodes of the edge are found to be conditionally independent. The threshold value for our TDPA implementation is 0.05.

For the experiments reported in the next section, we used the three previously described different subsets as such, as well as narrowed by means of using only the ten and twenty more commonly available variables in patients' clinical records. The rationale for this selection is, basically, trying to approximate the experiment to clinical reality constraints. The more often- recorded variables are likely to be the ones that are considered by the medical experts to be relevant in regular examinations at the ICU, for that reason, we focus on analyzing the predictive power of these most available inputs.

Finally, we also processed the graphs in the same situations but excluding the final outcome variable (deceased or not) from the data as to rise the hidden behaviours among the rest of measures that might have been eclipsed by the outcome variable. All the resulting figures can be seen in section 6: Results and Discussion.

Chapter 6

Methods for Classification

6.1 Support Vector Machines (SVM)

Support Vector Machine (SVM) is primarily a classier method that performs classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels. SVM supports both regression and classification tasks and can handle multiple continuous and categorical variables.

In this work we use SVM as a classification method. The training step involves the minimization of the following quadratic problem:

$$\frac{1}{2}w^T w + C\sum_{i=1}^{N}\zeta_i \tag{6.1}$$

subject to the constraints

$$y_i\big(w^T\phi(x_i) + b\big) \geq 1 - \zeta_i \tag{6.2}$$

and

$$\zeta_i \geq 0, i = 1, ..., N \tag{6.3}$$

where C is the capacity constant, w is the vector of coefficients, b is a constant, and represents parameters for handling nonseparable data (inputs). The index i labels the N training cases. Note that represents the class labels and x_i represents the independent variables. The kernel ϕ is used to transform data from the input (independent) to the feature space. It should be noted that the larger the C, the more the error is penalized. Thus, C should be chosen with care to avoid over fitting.

In this work we use the Gaussian Radial Basis Function (RBF) kernel, widely used in classification problems [78], which applies the following transformation:

$$\phi(\mathbf{x}, \mathbf{x'}) = exp\Big(-\frac{\|\mathbf{x} - \mathbf{x'}\|^2}{2\sigma^2}\Big) \tag{6.4}$$

6.2 Artificial Neural Networks (ANN)

Artificial Neural Networks (ANN) have been largely used as input-output mapping for different applications including modelling and classification [79]. The main characteristics of a neural network are parallel distributed structure and ability to learn, which produce excellent output for inputs not encountered during training. Moreover, the structure can be set to be simple enough to compute the output (or outputs) from the given inputs in very low computational time.

The basic processing elements of ANN are called artificial neurons, or simply neurons or nodes. Each processing unit is characterized by an activity level (representing the state of polarization of a neuron), an output value (representing the firing rate of the neuron), a set of input connections (representing synapses on the cell and its dendrites), a bias value (representing an internal resting level of the neuron), and a set of output connections (representing a neuron's axonal projections).

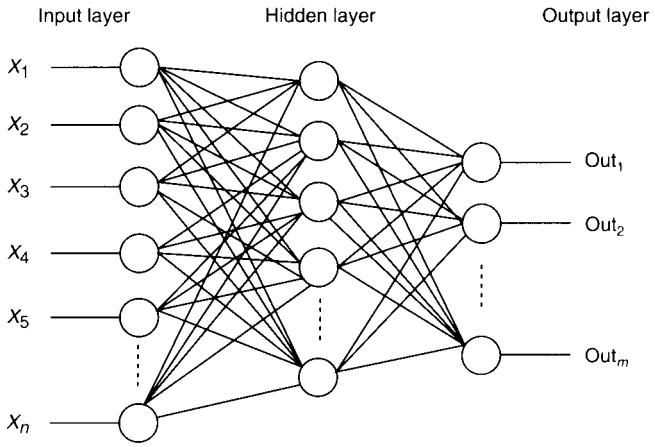


Figure 6.1: Typical Artificial Neural Network structure, with fully connected neurons.

The processing units are arranged in layers. There are typically three parts in a neural network: an input layer with units representing the input variables, one or more hidden layers, and an output layer with one or more units representing the output variables(s) (see figure 6.1). The units are joined with varying connections or weights. Each connection has an associated weight (synaptic strength) which determines the effect of the incoming input on the activation level of the unit. The weights may be positive (excitatory) or negative (inhibitory). The neuron output signal is given by the following relationship:

$$\sigma = f(\mathbf{w}^T x) = f\left(\sum_{j=1}^{n}(w_j x_j)\right) \tag{6.5}$$

where $\mathbf{w} = (w_1, ..., w_m)^T \in \mathbb{R}^n$ is the weight vector, and $x = (x_1, ..., x_n) \in \mathbb{R}^n$ is the vector of neuron inputs. The function $f(\mathbf{w}^T x)$ is often referred to as the activation (or transfer) function. Its domain is the set of activation values, net, of the neuron model, and is often represented by $f(net)$. The variable net is defined as a scalar product of the weight and input vectors:

$$net = \sum_{j=1}^{n}(w_j x_j) = w_1 x_1 + ... + w_n x_n \tag{6.6}$$

Training a neural network can be defined as the process of setting the weights of each connection between units in such a way that the network best fits to the underlying unknown function, thus turning it into an optimization problem. In this work is used *Levenberg-Marquardt* optimization method.

Chapter 7

Results and Discussion

7.1 Feature Selection for Classifiers

In this section we define the selected variables that were used in the two kind of the trained classifiers (SVM and ANN) for both outcome of the patient and SOFA prediction.

This phase is mainly driven by the information extracted from the application of Conditional Independence Maps (see section 7.2).

The first hypotheses for selecting the variables that are are going to be finally used in our classifiers is that they have to be easily measured in all ICU environments, in order to give meaning and applicability to the real world.

The variables used in designing the patient outcome predictor, the selected measurements were the following ones:

- Apache2
- MODS
- SAPS2
- SOFA
- Anticoagulation
- Analgesics
- SIRS
- PH
- Mean Blood Pressure
- CRP
- Temperature
- Systolic Blood Pressure
- Diastolic Blood Pressure
- O_2 Saturation

The dimension of the input in outcome prediction is of 14 variables, selected from the 103 measurements that compound MEDAN database.

In the case of predicting the Organ Dysfunction severity prediction, we decided to restricted even more the selected variables and to keep just the ones that are measured routinely and that are not directly related with Organ Dysfunction indicators, just with the exception of the C-Reactive Protein (CRP). Moreover, we made the experiments using just data coming from the first day (F1).

- Apache2
- Anticoagulation
- Analgesics
- SIRS
- PH
- Mean Blood Pressure
- CRP
- Temperature
- Systolic Blood Pressure
- Diastolic Blood Pressure
- O_2 Saturation

The size of Organ Dysfunction predictor subset is even smaller, consisting of 11 measurements. The decision of this variable subset is extracted from Causal Independence Maps experiments (see Chapter 5), using the variables that both are frequently and easily measured and are directly related with SOFA score.

As explained previously, SOFA score measures the dysfunction of six subsystems (respiratory, coagulation, liver, cardiovascular, central nervous system, and renal). A SOFA score greater than 7 implies a severe Organic Failure, so we use this statement to detect just admitted patients that presents a real risk to suffer that failure, so the clinicians can start facing with the potential new clinical picture.

7.2 Causal Discovery in Sepsis

In this section we describe the main results of applying Causal Independence Maps to our separated subsets, in order to bring hidden causal relationships to light.

The results obtained in our experiments, according to what was previously stated in the description of the experimental settings, are separated here into three different scenarios: Those corresponding to the first register of the time of admission in the ICU environment, those corresponding to the situation 24h later and, finally, those corresponding to the situation 48h from admission. In this way, it is possible to appreciate the differences in the roles that variables play at each stage of ICU stays motivated by sepsis. Also, experiments are separated according to whether the graphs include the outcome binary variable of death or not.

The Causal Independence Maps created by Causal Explorer for the experiments with the outcome variable are reported in figures 7.1 and 7.2, whereas those for the experiments without this outcome variable are reported in figures 7.3, 7.4 and 7.5.

An additional Conditional Independence Maps were created (can be found in Appendix 9 using not the tenth most dense variables but the twentieth ones, shown in figures A.1 and A.2. In the same way as above, we repeated the experiments excluding the outcome of the patients to enhance behaviours eclipsed by the attraction power of that variable, resulting in figures placed in appendix 9 (A.3, A.4 and A.5).

Discussion

The results displayed in Figures 7.1 and 7.2 throw some interesting light on the potential causal relationship between the analyzed variables and the outcome "death" variable. In particular, and interestingly, they are consistent and at least partially validate the official brand new modifications in the definition of sepsis [9] in two ways:

- The causal relationships found in the moment of admission in ICU and the scenario when 48 hours have elapsed are identical, so it can be concluded that the peak of activity in the development of the patients' situation takes place around 24 hours after being admitted at the ICU.
- The former definition of sepsis placed an excessive focus on inflammation and an inadequate specificity and sensitivity of the SIRS criteria. The scenario at 24h after admission ((and for this reason, a single figure is shown for both: see Figure 7.2 and A.2) neatly decouples SIRS from direct relationship with the final outcome, and it instead relates it to the Hemoglobin index.

Different time scenarios

1. *Admission time and 48h after admission time:* The graph in Figure 7.1 shows the situation of causal relationships in the admission moment, specifically the behaviour of the 10 most recorded variables in the data set (the ones with the lowest proportion of missing values). The results for two days later from admission were exactly identical.

   The links between SIRS and the outcome variable at these times are remarkable and maintain a two-way connection, meaning a strong relationship among the SIRS diagnosis and the chances of the patient to decease. The rest of variables include several physiological measurements (Heart Rate, Systolic Blood Pressure, etc.) and standard ICU calculated scores such as SOFA and Apache II. The results with the 20 variables with the lowest proportion

of missing values, reported in Figure A.1, again show a very strong link between SIRS and the outcome and only Analgesics does not relate directly with "death", but through the Hematocrit measurement.

2. *24h after admission:* Figures 7.2 for 10 variables (as before) and A.2 for 20 variables would be identical to the ones at admission time and 48h after admission were it not for the already mentioned decoupling of SIRS from the outcome variable. It is worth noting that, in this case, SIRS indirectly related to the outcome through Hematocrit (a test indicating the percentage of the volume of whole blood that is made up of red blood cells) for 20 variables, and through Hemoglobin for 10 variables. It has been reported [80]that in the acute phase of sepsis, several potential mechanisms may change the Hemoglobin concentration in the blood. Analgesics is this time indirectly related to the outcome through Creatinine Serum instead of Hematocrit.

Maps excluding the outcome variable

A detailed analysis of the relationships between the measured variables when the outcome variable is removed from the analysis is beyond the remit of the current brief paper. It is nevertheless worth mentioning that, again, maps at admission and after 48h are identical and both different from the maps at 24h after admission. It is also noteworthy that, at admission and after 48h, SIRS is quite isolated from the rest of variables, relating to them only through the SOFA core for 10 variables and through Hematocrit for 20 variables, whereas for data at 24h after admission, many variables have a direct impact on SIRS.

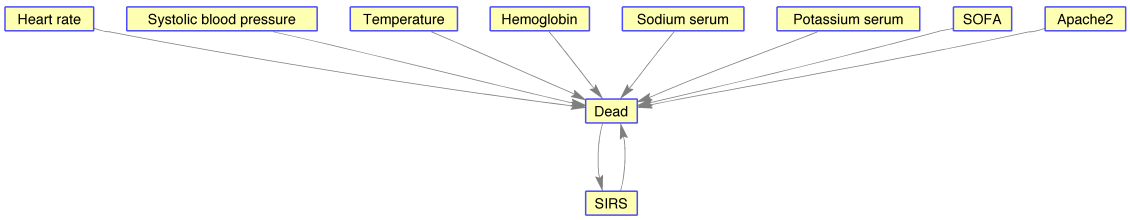


Figure 7.1: Conditional Independence Map for data at the time of admission and 48h after admission, using the 10 variables with lowest missingness ratio.

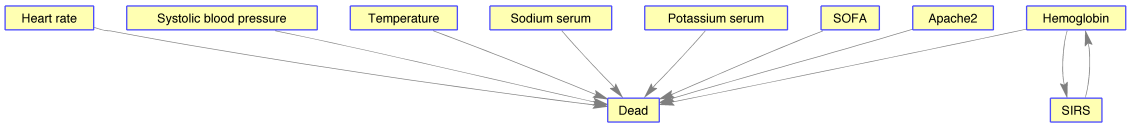


Figure 7.2: Conditional Independence Map for data 24h after admission, using the 10 variables with lowest missingness ratio.

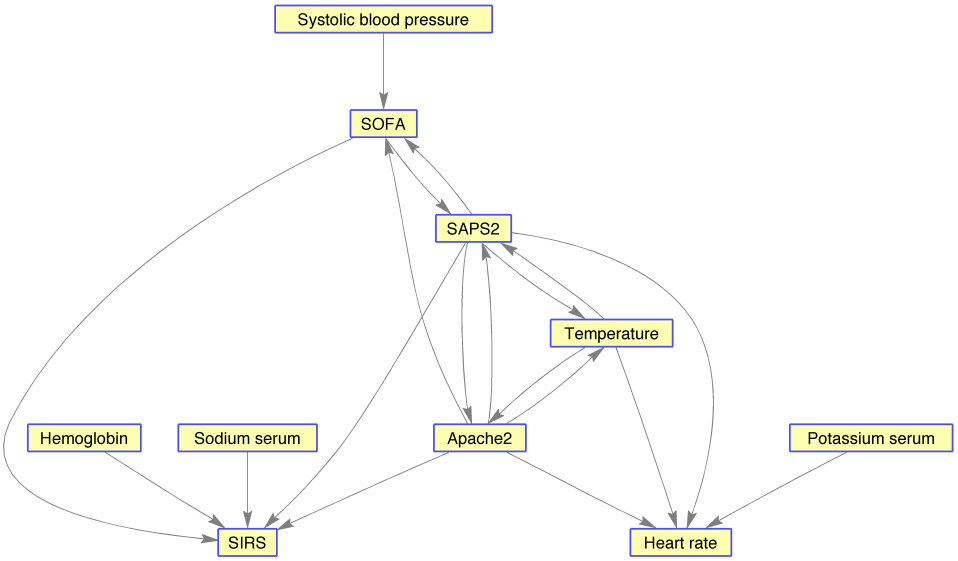


Figure 7.3: Conditional Independence Map for data at the moment of admission, using the 10 variables with lowest missingness ratio.

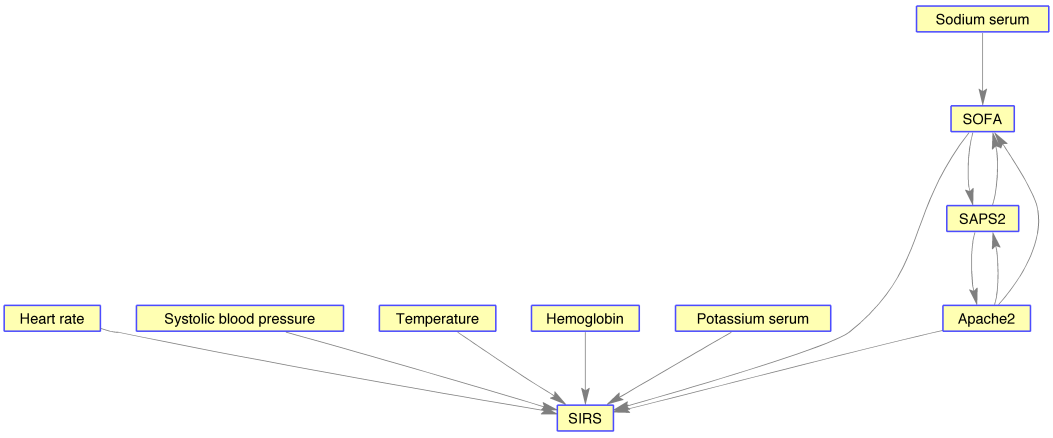


Figure 7.4: Conditional Independence Map for data 24h after admission, using the 10 variables with lowest missingness ratio.

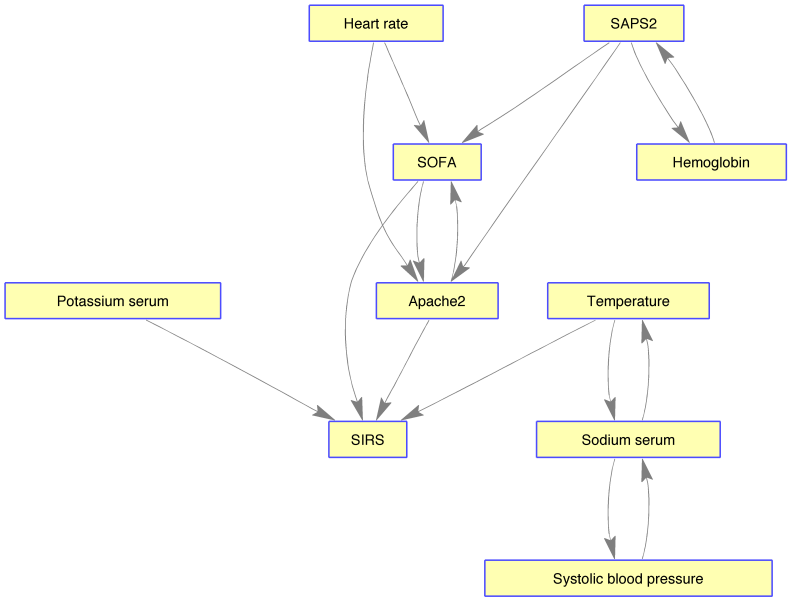


Figure 7.5: Conditional Independence Map for data 48h after admission, using the 10 variables with lowest missingness ratio.

7.3 Death Prediction in Patients

In this section we present the results of applying the classification techniques detailed in chapter 6.

One of the objective of this work is to find significant improvements in early diagnoses in the first evaluation of the patient, using information from just their admission in ICU and, at most, one or two days later, since it does not make any sense to design a very nice predictor that is based in patient measurements taken just moments before his *Sepsis* or aggravation episode.

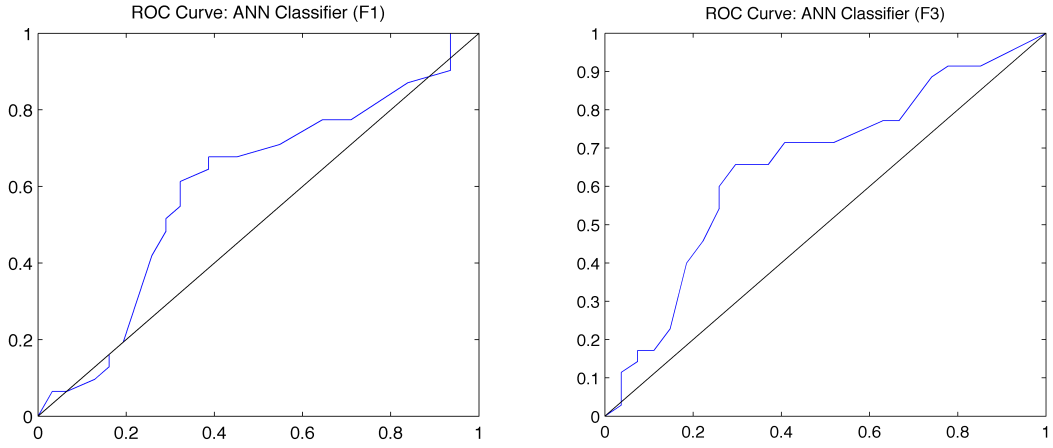
Brause *et. al.*, the authors of the MEDAN database, published an article [81] with their results of the MEDAN Project, and they concluded that with the data gathered during the first three days (F3) of admission in ICU, it is not possible to get a reliable diagnosis. In the cited article they got an area beneath the ROC curve (AUC) of AUC=0.56, value that indicates almost randomly results by their trained Neural Network. As could be expected, they found the best results considering the data of the last day, with AUC=0.93, but it is not useful at all for building an alarm system. So we applied the two classification techniques described in this work (Artificial Neural Network and Support Vector Machines) to the "time of admission" subset (F1), and the results are detailed in table 7.1.

	ANN		**SVM**	
	AUC	*Accuracy*	*AUC*	*Accuracy*
Brause et al (using F3)	0.56	-	-	-
Using F1	0.59	0.56	0.54	0.55
Using F3	0.66	0.59	0.62	0.55

Table 7.1: Results of ANN classifier using F1 (first day of admission data).

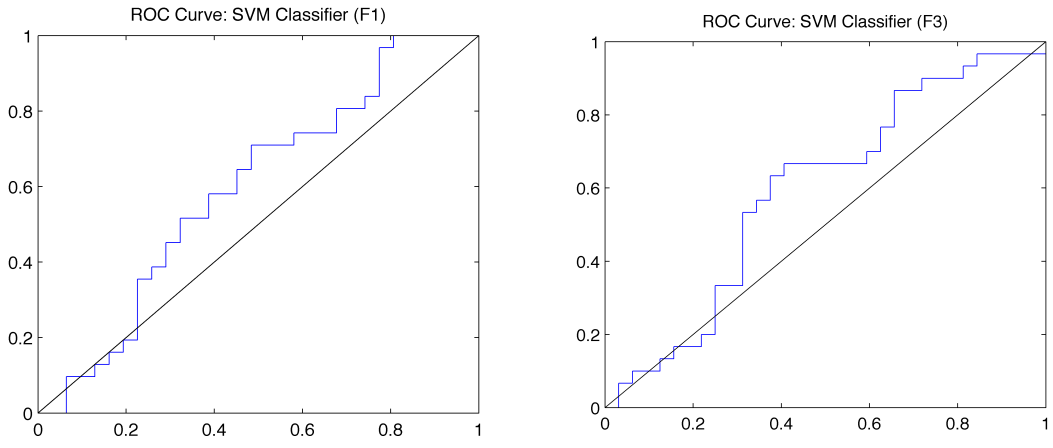
Neural Network Classifier

The procedure in designing the binary classifier by means of an Artificial Neural Network was in this way:



(a) *(b)*

Figure 7.6: ROC curve using ANN classifier, using F1 (a) and F3 (b) data set.



(a) *(b)*

Figure 7.7: ROC curve using SVM classifier, using F1 (a) and F3 (b) data set.

- Select the data to use (F1 or F3).
- Split the data set into train and test subsets, with ratio 85/15 respectively.
- Train networks with different number of neurons in the hidden layer (from 5 to 60).
- Evaluate the performance in test subset.

In the case of the data gathered the first day of admission (F1), since the size of the input is smaller than taking into account the first three days (F3), the optimal number of artificial neurons in the hidden layer is of 30. In the other hand, when designing the network to predict with data coming from the first three days (F3) the optimal number of hidden units is about 50.

Support Vector Machines Classifier

The SVM classifier used in this section is a standard one, using the Radial Basis Function kernel, widely used in classification tasks, and specially in clinical data [42]. The cross validation step is done by 10-fold procedure, and the train/test ratio is defined to 85/15; in the same way than the ANN experiments.

- Select the data to use (F1 or F3).
- Split the data set into train and test subsets, with ratio 85/15 respectively.
- Apply 10-fold Cross Validation in training set.
- Evaluate the resulting classifier in test set.

7.4 Organ Dysfunction Prediction in Patients

In the previous section we made a pure predictor of the outcome of the patient and we found better results than a previous similar study from Brause *et al* [81], but still not good enough results to state that with information of the patient coming from their first hours in ICU is possible to reliably predict their chance to end up deceased.

Given that the main aim of this work is to help clinicians in early detection and prediction of possible complications in patients status, in order to prevent them of deceasing, we decided to look for another indicator, apart from the proper "death" output, as to give physicians reliable information about the potential severity of the patient.

Hence, we decided to use SOFA to describe Organ Dysfunction and failure, as many works support ([41]), separating the patients having a SOFA score greater than 7 into the potential population of suffering from severe consequences in the following hours, and some of them eventually die.

	ANN		**SVM**	
	AUC	*Accuracy*	*AUC*	*Accuracy*
Using F1	0.74	0.75	0.77	0.72

Table 7.2: Results of applying SVM and ANN classifier to predict Severe Organ Failure episodes, using measures of the admission moment in ICU.

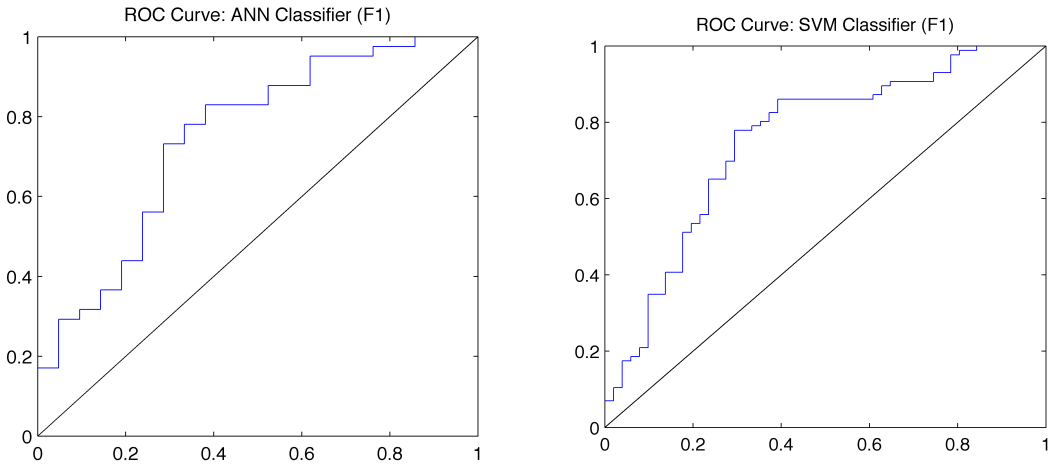
7.5 Results

The results turn out to be much better than the direct "death" prediction (achieving accuracies of 75%, versus the 56% in outcome prediction in the same conditions) , and even we predict the "gravity" and not the final of the patient, it gives valuable information to medical crew to start the procedures that they believe necessary.

The procedures of SVM and ANN design and training are exactly the same than the previous section of outcome prediction.

The data used in the organ dysfunction classifiers are the first recordings in the admission in ICU (F1), so it gives an important added value in the applicability in the constraints that an ICU environment requires.

The Area Under ROC Curve is similar in the two used models (0.74 in ANN and 0.77 in SVM), so it shows a clear advance with respect to predicting directly the outcome of the patient.



(a) *(b)*

Figure 7.8: ROC curve using ANN (a) and SVM (b) Organ Failure classifier, using F1 data set.

Chapter 8

Conclusions and Future Work

In the previous chapters, we have first defined the general problem of *Sepsis* data analysis in the ICU environment and we have then focused our attention on some of the main challenges it involves, including the study of the causal relationships among the most recorded variables, the prediction of ICU outcome for patients with *Severe Sepsis* and an organ dysfunction assessment.

The Systemic Inflammatory Response Syndrome (SIRS) assigned role within the *Sepsis* definition has kept changing over the last decades, partly due to the inexact definition of *Sepsis*; it is in this context that this thesis tries to make some contribution.

In the causal relationships study in chapter 5, we have seen the different interrelated behaviours of variables. The comparisons of the results in the moment of admission in ICU versus the situation 24 and 48h later yield encouragingly good results and they seem to support the new definition of *Sepsis* [9], which is placing more emphasis in organ dysfunction than inflammation (and, therefore, SIRS). In our results, the SIRS variable is clearly decoupled from direct relationship with the final outcome (i.e. exitus letalis in the ICU).

However, a word of caution must be given since our results also support an important role of SIRS in organ dysfunction measured by the SOFA score. This comes as no surprise since the relation of inflammation and organ dysfunction since it is the former that mediates the latter. Organ dysfunction by itself is not specific of sepsis and this is clearly shown in the graphs that we obtained taking organ dysfunction as the variable of study instead of ICU outcome.

From our results one may well conclude that it is the organ dysfunction that drives the ICU outcome but also it is SIRS that drives the organ dysfunction in Sepsis. For this reason we believe that the definition of sepsis should be revisited once again.

The Causal Discovery algorithms applied to septic patients gave also valuable information for the feature selection stage in the design of an outcome predictor, as we used the direct relationships to our response variable to create a tight variable subset with the most recorded variables.

Finally, we based the design of the classification system for both assessing ICU outcome and organ dysfunction in the information discovered in the previous sections, and the obtained results were rather promissing.

In the first case, when predicting the outcome of the patient, our results showed identical behaviours than those reported by Paetz *et al.* [5], even though they used data extracted from the first three days of hospitalization in ICU, while we used just the data at admission (i.e. in the phase of most acuity). In this regard, our results are better since we achieve the same accuracy at an earlier stage. This of course may show better prognosis since clinical protocols show that administration of antibiotics during the first 6 hours result in much better ICU outcomes.

The second classifier experiment carried out concerned the detection of whether a patient presented a clinical picture that could lead to severe Organ Dysfunction and, therefore, a serious risk of death. In this case, even though we are not predicting the proper outcome of the patient, it even more important because it provides clinicians with valuable information to begin dealing with this potential threat at an early stage.

The results of this experiment were also promising, specially taking into account that we just used data of the admission at the ICU. For both the Artificial Neural Network and the Support Vector Machine based classifiers, the accuracy was in the area of 75%.

In conclusion, the experiments presented in this work are focused to both improve the understandability of *Sepsis* and its prognosis, with the help of Artificial Intelligence approaches, while, importantly, at the same time respecting the ICU constraints and needs of rapid action. Our results reinforce the initial decision to use data acquired in the first moments of ICU stay in accordance with standard clinical practice.

8.0.1 Outline for Future Work

One of the main contributions of this master thesis is the study of the usefulness of SIRS in the continuum of *Sepsis*, by means of the application of Conditional Independence Maps, and we got promising results according to new definition of *Sepsis*, but the potential power of these graphs can far contribute to new advances in the yet poor understanding of what causes *Sepsis*. We believe that the debate about using SIRS for the definition of sepsis is not yet closed.

Our work focused in making decisions at ICU admission and then during the first three days of evolution. Of course, these are the most critical moments in the continuum of sepsis. However, it may also be important to put this information in a more dynamic context. Taking the time evolution of sepsis with graphical models is a very interesting problem that could be tackled as future work combining clinical traits such as those used in this thesis with data more specifically related to inflammation and organ dysfunction such as proteins and metabolites. This is a very active area of research and we believe that the methods presented here could contribute to providing more insight about the pathophysiology of sepsis and even set a methodological basis for finding potential biomarkers of sepsis.

Chapter 9

Thesis publications

Beyond the work reported in this thesis, the study of Causal Relationships among patients' variables during their admission and first hours at the ICU led to the publication of a conference paper, whose details are provided below.

Applying Conditional Independence Maps to improve Sepsis Prognosis

Conference: Data Mining in Biomedical Informatics and Healthcare (DMBIH) Workshop, part of the IEEE International Conference on Data Mining (ICDM 2016), Barcelona, Spain
Date of publication: December 12th, 2016
Authors: Carles Morales, Vicent Ribas, Alfredo Vellido.

Abstract — Sepsis has become a major challenge to medicine and day-to-day clinical practice due to its prevalence in the Inten- sive Care Unit. This transversal condition has high prevalence and considerable risk of death in its most developed stages. Sepsis has recently been officially redefined and an important new trait of this updated definition is that organ dysfunction is now taken into account, replacing the focus on Systemic Inflammatory Response Syndrome of the previous definition. In this brief study, we analyze one of the biggest available multi-centre sepsis databases using Conditional Independence Maps methods. With this, we aim to explore potential causal relationships between the measured variables and the survival outcome and also to validate the changes in the new definition of sepsis.
keywords— *Sepsis, Conditional Independence Maps, Intensive Care Unit.*

Appendix A

Causal Independence Maps

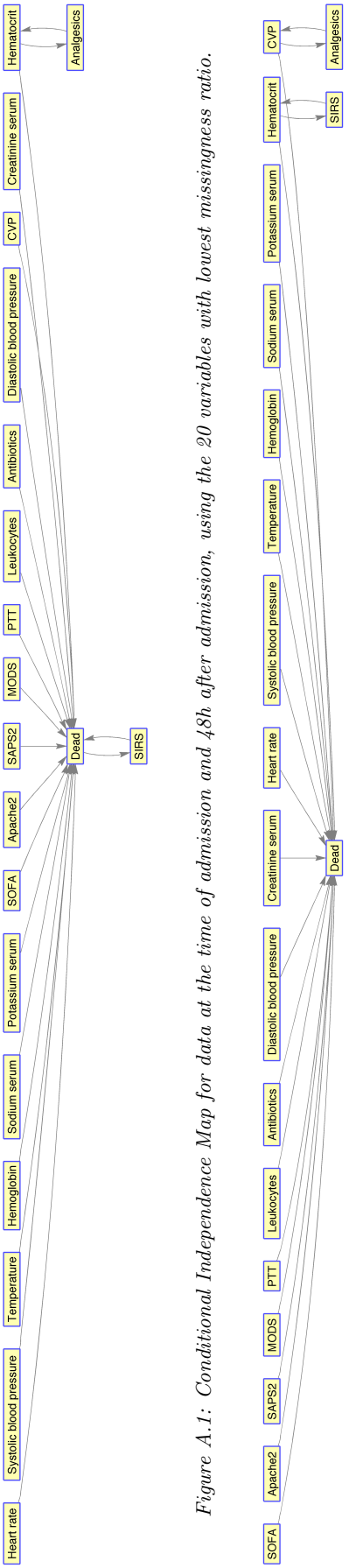


Figure A.1: Conditional Independence Map for data at the time of admission and 48h after admission, using the 20 variables with lowest missingness ratio.

Figure A.2: Conditional Independence Map for data 24h after admission, using the 20 variables with lowest missingness ratio.

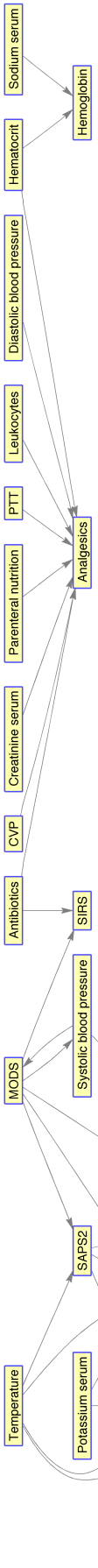


Figure A.3: Conditional Independence Map for data at the time of admission, using the 20 variables with lowest missingness ratio.

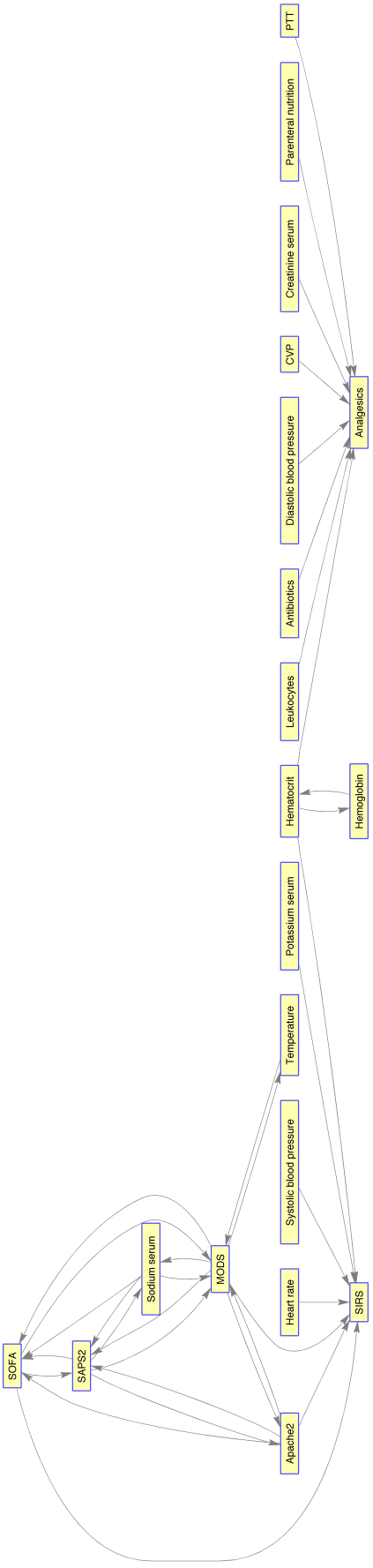


Figure A.4: Conditional Independence Map for data 24h after admission, using the 20 variables with lowest missingness ratio.

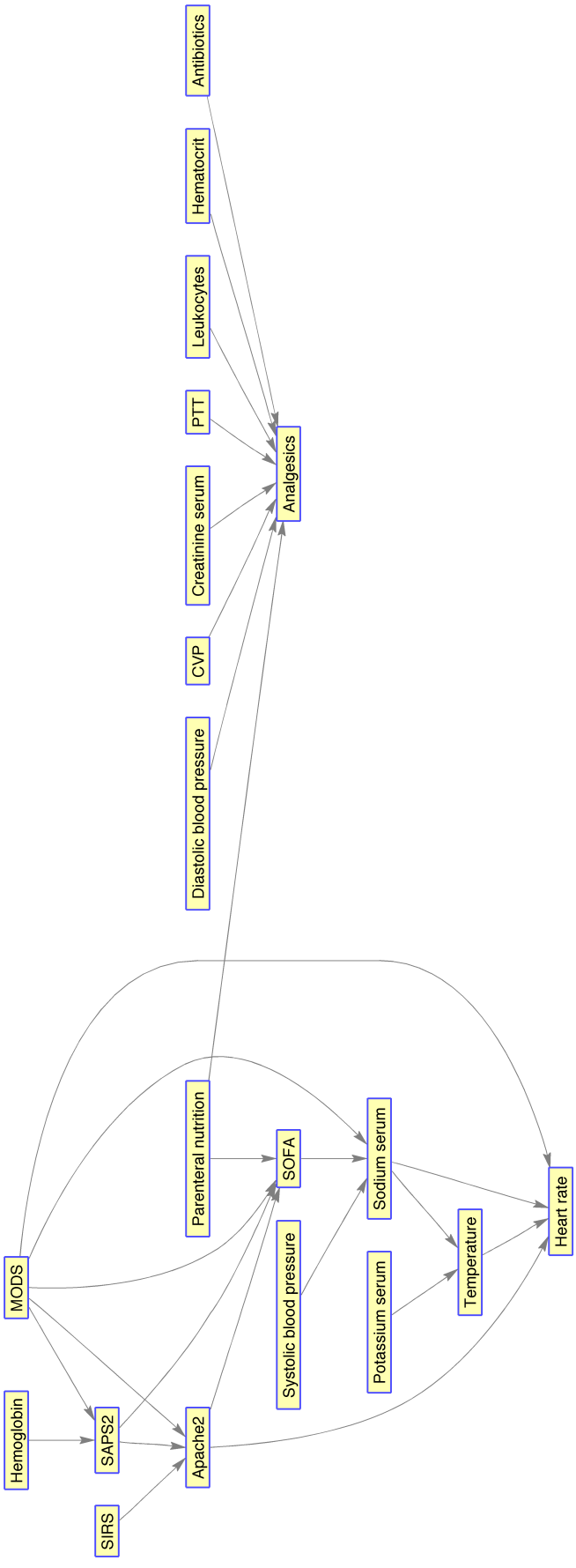


Figure A.5: Conditional Independence Map for data 48h after admission, using the 20 variables with lowest missingness ratio.

References

[1] Pajman Danai and Greg S Martin. "Epidemiology of sepsis: recent advances". In: *Current infectious disease reports* 7.5 (2005), pp. 329–334.

[2] Niranjan Kissoon et al. "World Federation of Pediatric Intensive Care and Critical Care Societies: Global Sepsis Initiative*". In: *Pediatric Critical Care Medicine* 12.5 (2011), pp. 494–503.

[3] Derek C Angus et al. "Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care". In: *Critical Care Medicine-Baltimore-* 29.7 (2001), pp. 1303–1310.

[4] Corinne Alberti et al. "Epidemiology of sepsis and infection in ICU patients from an international multicentre cohort study". In: *Intensive care medicine* 28.2 (2002), pp. 108–121.

[5] Jürgen Paetz and Björn Arlt. "A neuro-fuzzy based alarm system for septic shock patients with a comparison to medical scores". In: *Medical Data Analysis*. Springer, 2002, pp. 42–52.

[6] Greg S Martin et al. "The epidemiology of sepsis in the United States from 1979 through 2000". In: *New England Journal of Medicine* 348.16 (2003), pp. 1546–1554.

[7] H Shaw Warren. "Editorial: Mouse models to study sepsis syndrome in humans". In: *Journal of leukocyte biology* 86.2 (2009), pp. 199–201.

[8] Emanuel Rivers et al. "Early goal-directed therapy in the treatment of severe sepsis and septic shock". In: *New England Journal of Medicine* 345.19 (2001), pp. 1368–1377.

[9] Mervyn Singer et al. "The third international consensus definitions for sepsis and septic shock (sepsis-3)". In: *Jama* 315.8 (2016), pp. 801–810.

[10] J Paetz et al. "[The MEDAN database: patients with abdominal septic shock]". In: *Zentralblatt fur Chirurgie* 128.4 (2003), pp. 298–303.

[11] Jürgen Paetz et al. "Data quality aspects of a database for abdominal septic shock patients". In: *Computer methods and programs in biomedicine* 75.1 (2004), pp. 23–30.

[12] R Brause et al. "Neural Networks for sepsis prediction-the MEDAN-Project". In: *JOURNAL FUR ANASTHESIE UND INTENSIVBEHANDLUNG* 11 (2004), pp. 40–43.

[13] André S Fialho et al. "Predicting outcomes of septic shock patients using feature selection based on soft computing techniques". In: *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer. 2010, pp. 65–74.

[14] Rúben DMA Pereira et al. "Predicting septic shock outcomes in a database with missing data using fuzzy modeling: Influence of pre-processing techniques on real-world data-based classification". In: *Fuzzy Systems (FUZZ), 2011 IEEE International Conference on*. IEEE. 2011, pp. 2507–2512.

[15] Theodore J Iwashyna et al. "Identifying patients with severe sepsis using administrative claims: patient-level validation of the angus implementation of the international consensus conference definition of severe sepsis". In: *Medical care* 52.6 (2014), e39.

[16] Roger C Bone. "Sepsis, the sepsis syndrome, multi-organ failure: a plea for comparable definitions". In: *Annals of internal medicine* 114.4 (1991), pp. 332–333.

[17] Roger C Bone. "The pathogenesis of sepsis". In: *Annals of Internal Medicine* 115.6 (1991), pp. 457–469.

[18] David JJ Muckart and Satish Bhagwanjee. "American College of Chest Physicians / Society of Critical Care Medicine Consensus Conference definitions of the systemic inflammatory response syndrome and allied disorders in relation to critically injured patients". In: *Critical care medicine* 25.11 (1997), pp. 1789–1795.

[19] Paul E Marik and Joseph Varon. "Sepsis: state of the art". In: *Disease-A-Month* 47.10 (2001), pp. 462–532.

[20] Patrik Gille-Johnson. *Diagnostic and prognostic markers in sepsis*. Inst för medicin, Huddinge/Dept of Medicine, Huddinge, 2013.

[21] FM Brunkhorst et al. "Epidemiology of severe sepsis and septic shock in Germany: results from the German'Prevalence'study". In: *Critical Care* 9 (2005), pp. 1–2.

[22] Christopher W Seymour et al. "Severe sepsis in pre-hospital emergency care: analysis of incidence, care, and outcome". In: *American journal of respiratory and critical care medicine* 186.12 (2012), pp. 1264–1271.

[23] Maria del Rocio Perez Benitez. "Actualización de la sepsis en adultos. Código sepsis". PhD thesis. Universidad Internacional de Andalucía, 2015.

[24] Greg S Martin. "Sepsis, severe sepsis and septic shock: changes in incidence, pathogens and outcomes". In: *Expert review of anti-infective therapy* 10.6 (2012), pp. 701–706.

[25] James C Jackson et al. "Acute respiratory distress syndrome, sepsis, and cognitive decline: a review and case study". In: *Southern medical journal* 102.11 (2009), p. 1150.

[26] C Niels, Ren-Feng Guo Riedemann, GUO FENG, et al. "The enigma of sepsis". In: *J Clin Invest* 112 (2003), pp. 460–467.

[27] Lewis Thomas. "Germs". In: *New England Journal of Medicine* 287.11 (1972), pp. 553–555.

[28] RCea Bone et al. "American-College of Chest Physicians Society of Critical Care Medicine Consensus Conference-definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis". In: *Critical care medicine* 20.6 (1992), pp. 864–874.

[29] Mitchell M Levy et al. "2001 sccm/esicm/accp/ats/sis international sepsis definitions conference". In: *Intensive care medicine* 29.4 (2003), pp. 530–538.

[30] Michael D Howell et al. "Proof of principle: The predisposition, infection, response, organ failure sepsis staging system*". In: *Critical care medicine* 39.2 (2011), pp. 322–327.

[31] Roger C Bone et al. "Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis". In: *Chest* 101.6 (1992), pp. 1644–1655.

[32] MP Rodrigo Casanova et al. "Sepsis grave y shock séptico". In: *Rev. Esp. Anestesiol. Reanim* 54 (2007), pp. 484–498.

[33] Bernhard Schlüter et al. "Effect of the interleukin-6 promoter polymorphism (- 174 G/C) on the incidence and outcome of sepsis". In: *Critical care medicine* 30.1 (2002), pp. 32–37.

[34] Christina Wacker et al. "Procalcitonin as a diagnostic marker for sepsis: a systematic review and meta-analysis". In: *The Lancet infectious diseases* 13.5 (2013), pp. 426–435.

[35] Pedro Póvoa. "C-reactive protein: a valuable marker of sepsis". In: *Intensive care medicine* 28.3 (2002), pp. 235–243.

[36] Abderr azzaq Belaaouaj, Kwang Sik Kim, and Steven D Shapiro. "Degradation of outer membrane protein A in Escherichia coli killing by neutrophil elastase". In: *Science* 289.5482 (2000), pp. 1185–1187.

[37] P Ricart and J Morillas. "Marcadores de respuesta inflamatoria sistémica". In: *Sepsis grave. Barcelona: Ediciones experiencia* (2002), pp. 59–65.

[38] Vicent J Ribas Ripoll and Enrique Romero Merino. "On the Intelligent Management of Sepsis in the Intensive Care Unit". In: (2012).

[39] Frank M Collins and Nancy N Pearsall. "Cellular antimicrobial immunity". In: *CRC critical reviews in microbiology* 7.1 (1979), pp. 27–91.

[40] Christopher C Baker and Toan Huynh. "Sepsis in the critically ill patient". In: *Current problems in Surgery* 32.12 (1995), pp. 1013–1083.

[41] J-L Vincent et al. "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure". In: *Intensive care medicine* 22.7 (1996), pp. 707–710.

[42] Vicent J Ribas et al. "Severe sepsis mortality prediction with logistic regression over latent factors". In: *Expert Systems with Applications* 39.2 (2012), pp. 1937–1943.

[43] Mitchell M Levy et al. "Early changes in organ function predict eventual survival in severe sepsis". In: *Critical care medicine* 33.10 (2005), pp. 2194–2201.

[44] Derek C Angus and Tom Van der Poll. "Severe sepsis and septic shock". In: *New England Journal of Medicine* 369.9 (2013), pp. 840–851.

[45] Stanley Lemeshow et al. "Mortality Probability Models (MPM II) based on an international cohort of intensive care unit patients". In: *Jama* 270.20 (1993), pp. 2478–2486.

[46] Jean-Roger Le Gall, Stanley Lemeshow, and Fabienne Saulnier. "A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study". In: *Jama* 270.24 (1993), pp. 2957–2963.

[47] Jean-Roger Le Gall et al. "A simplified acute physiology score for ICU patients." In: *Critical care medicine* 12.11 (1984), pp. 975–977.

[48] Prashant Prakash, Kavita Krishna, and Deepansh Bhatia. "Usefulness of SAPS II Scoring System as an Early Predictor of Outcome in ICU Patients". In: ().

[49] Jean Roger Le Gall et al. "Mortality prediction using SAPS II: an update for French intensive care units". In: *Critical Care* 9.6 (2005), p. 1.

[50] Rui P Moreno et al. "SAPS 3—From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission". In: *Intensive care medicine* 31.10 (2005), pp. 1345–1355.

[51] Vicent J Ribas et al. "On the use of decision trees for ICU outcome prediction in sepsis patients treated with statins". In: *Computational Intelligence and Data Mining (CIDM), 2011 IEEE Symposium on*. IEEE. 2011, pp. 37–43.

[52] Vicent J Ribas Ripoll et al. "Sepsis mortality prediction with the Quotient Basis Kernel". In: *Artificial intelligence in medicine* 61.1 (2014), pp. 45–52.

[53] William A Knaus et al. "APACHE II: a severity of disease classification system." In: *Critical care medicine* 13.10 (1985), pp. 818–829.

[54] LM Wier et al. "HCUP facts and figures: statistics on hospital-based care in the United States, 2009". In: *Rockville, MD: Agency for Healthcare Research and Quality* (2011).

[55] Anand Kumar et al. "Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock". In: *Critical care medicine* 34.6 (2006), pp. 1589–1596.

[56] Eren Gultepe et al. "From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system". In: *Journal of the American Medical Informatics Association* 21.2 (2014), pp. 315–325.

[57] Linda Peelen et al. "Using hierarchical dynamic Bayesian networks to investigate dynamics of organ failure in patients in the Intensive Care Unit". In: *Journal of biomedical informatics* 43.2 (2010), pp. 273–286.

[58] Rüdiger W Brause. "Medical analysis and diagnosis by neural networks". In: *International Symposium on Medical Data Analysis*. Springer. 2001, pp. 1–13.

[59] R Brause, F Hamker, and Jürgen Paetz. "Septic shock diagnosis by neural networks and rule based systems". In: *Computational intelligence processing in medical diagnosis*. Springer, 2002, pp. 323–356.

[60] Shu-Li Wang, Fan Wu, and Bo-Hang Wang. "Prediction of severe sepsis using SVM model". In: *Advances in computational biology*. Springer, 2010, pp. 75–81.

[61] Collin HH Tang et al. "Non-invasive classification of severe sepsis and systemic inflammatory response syndrome using a nonlinear support vector machine: a preliminary study". In: *Physiological measurement* 31.6 (2010), p. 775.

[62] Vicent Jordi Ribas Ripoll et al. "On the intelligent management of sepsis in the intensive care unit". In: (2013).

[63] Christopher J Langmead. "Generalized queries and bayesian statistical model checking in dynamic bayesian networks: Application to personalized medicine". In: (2009).

[64] Mehmet Kayaalp, Gregory F Cooper, and Gilles Clermont. "Predicting ICU mortality: a comparison of stationary and nonstationary temporal models." In: *Proceedings of the AMIA symposium*. American Medical Informatics Association. 2000, p. 418.

[65] Vicent J Ribas et al. "On the use of graphical models to study ICU outcome prediction in septic patients treated with statins". In: *International Meeting on Computational Intelligence Methods for Bioinformatics and Biostatistics*. Springer. 2011, pp. 98–111.

[66] Nathan I Shapiro et al. "Mortality in Emergency Department Sepsis (MEDS) score predicts 1-year mortality". In: *Critical care medicine* 35.1 (2007), pp. 192–198.

[67] Ewa Straszecka. "Combining uncertainty and imprecision in models of medical diagnosis". In: *Information Sciences* 176.20 (2006), pp. 3026–3059.

[68] Richard E Neapolitan. *Probabilistic methods for bioinformatics: with an introduction to Bayesian networks*. Morgan Kaufmann, 2009.

[69] Richard E Neapolitan et al. "Learning bayesian networks". In: (2004).

[70] Peter Spirtes, Clark N Glymour, and Richard Scheines. *Causation, prediction, and search*. MIT press, 2000.

[71] Clark N Glymour and Gregory Floyd Cooper. *Computation, causation, and discovery*. Aaai Press, 1999.

[72] FT De Dombal et al. "Computer-aided diagnosis of acute abdominal pain". In: *Br Med J* 2.5804 (1972), pp. 9–13.

[73] RE Neapolitan. "Probabilistic reasoning in expert systems: theory and applications". In: *New York: Whey* (1990).

[74] Constantin F Aliferis et al. "Causal Explorer: A Causal Probabilistic Network Learning Toolkit for Biomedical Discovery." In: *METMBS*. Vol. 3. 2003, pp. 371–376.

[75] Judea Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 2014.

[76] Finn V Jensen. *An introduction to Bayesian networks*. Vol. 210. UCL press London, 1996.

[77] Gregory F Cooper and Edward Herskovits. "A Bayesian method for the induction of probabilistic networks from data". In: *Machine learning* 9.4 (1992), pp. 309–347.

[78] Mohamad T Musavi et al. "On the training of radial basis function classifiers". In: *Neural networks* 5.4 (1992), pp. 595–603.

[79] Simon S Haykin et al. *Neural networks and learning machines*. Vol. 3. Pearson Upper Saddle River, NJ, USA: 2009.

[80] G Jansma et al. "Is hemoglobin concentration affected by sepsis in the acute phase?" In: *Critical Care* 17.2 (2013), p. 1.

[81] R Brause et al. "Predicting Death for Abdominal Septic Shock Patients-The Results of the MEDAN Project". In: *Nierhaus, KG Kreymann (Eds.): Sepsis, SIRS, Immune Response-Concepts, Diagnostics and Therapy, Pabst Science Publ., Lengerich* (2005), pp. 119–126.